621.39
С56

Б. Я. СОВЕТОВ

# ТЕОРИЯ ИНФОРМАЦИИ

## ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПЕРЕДАЧИ ИНФОРМАЦИИ В АСУ

621.39
С56

Б. Я. СОВЕТОВ

# ТЕОРИЯ ИНФОРМАЦИИ

## ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПЕРЕДАЧИ ИНФОРМАЦИИ В АСУ

Б. Я. СОВЕТОВ ТЕОРИЯ ИНФОРМАЦИИ

068805

621.39
С56

МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО
СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ РСФСР

*Б. Я. СОВЕТОВ*

# ТЕОРИЯ ИНФОРМАЦИИ

## Теоретические основы передачи информации в АСУ

Допущено Министерством высшего и среднего
специального образования СССР
в качестве учебного пособия
для студентов вузов, обучающихся по специальности
автоматизированные системы управления

068805
621.391.23

ИЗДАТЕЛЬСТВО ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА
ЛЕНИНГРАД 1977

2013
ПЕРЕВІРЕНО

*Рекомендовано к изданию Головным советом по системам управления и средствам автоматики*

*УДК 681.323*

**Советов Б. Я. Теория информации.** Л., Изд-во Ленингр. ун-та, 1977. 184 с. Ил.—79, библиогр.—30 назв.

В учебном пособии изложены теоретические основы передачи информации в автоматизированных системах управления (АСУ), рассмотрены содержание и особенности автоматизированного управления производством, состав комплекса технических средств АСУ, общие положения об информационном процессе в АСУ. Вопросы передачи информации проанализированы на базе фундаментальных положений теории информации о количественной мере информации, соотношении скорости передачи дискретной и непрерывной информации и пропускной способности канала связи. Значительное внимание уделено передаче информации по каналам связи без шума и с шумом, информационным пределам избыточности, методам передачи непрерывной информации, способам выделения сигналов на фоне помех, а также информационному подходу к оценке качества функционирования комплекса технических средств АСУ.

Пособие предназначается для студентов, специализирующихся в области АСУ, а также может быть полезно для инженерно-технических работников, занимающихся разработкой технического и информационного обеспечения АСУ.

Рецензенты: кафедра информационных и управляющих систем Ленинградского политехнического ин-та им. М. И. Калинина; доктор техн. наук проф. В. И. Николаев (СЗПИ).

С $\frac{30\,501-123}{076(02)-77}$ 132—77

© Издательство Ленинградского университета, 1977 г.

ИБ № 439

*Борис Яковлевич Советов*

**Теория информации**
(теоретические основы передачи информации в АСУ)

Редакторы *Ф. И. Шаренкова,*
*Г. И. Чередниченко*
Техн. редактор *А. В. Борщева*
Корректоры *С. К. Школьникова, А. С. Качинская*

Сдано в набор 21/II 1977 г. Подписано к печати 7/IX 1977 г. Формат бумаги 60×90 1/16. Бум. тип. № 1. Печ. л. 11,5. Уч.-изд. л. 10,72. Тираж 13605 экз. Заказ 931. Цена 53 коп.
Издательство ЛГУ им. А. А. Жданова. 199 164 Университетская наб., 7/9

Типография им. Анохина
Управления по делам издательств, полиграфии и книжной торговли
Совета Министров Карельской АССР,
Петрозаводск, ул. «Правды», 4.

## ПРЕДИСЛОВИЕ

Повышение эффективности управления производством невозможно без внедрения автоматизированных систем управления (АСУ), существенной особенностью которых является присутствие человека в контуре управления. В АСУ в первую очередь автоматизируются процессы обработки, сбора, передачи, хранения, отображения и регистрации информации, а поэтому наряду с разработкой организационного и функционального аспектов огромное значение приобретает технический аспект АСУ, отображающий комплекс технических средств. Изучая информационный процесс, протекающий в АСУ, можно выделить в комплексе технических средств следующие подсистемы: подготовки и регистрации информации, сбора и передачи информации, хранения и обработки информации, выдачи и воспроизведения информации. В каждой из подсистем осуществляется преобразование информации с заданными показателями качества.

Общие закономерности преобразования информации изучает теория информации. В ряду таких дисциплин подготовки инженеров по АСУ, как „Исследование операций", „Основы построения АСУ" и др., „Теория информации" занимает базовое положение, определяя основные закономерности протекания информационного процесса. Теория информации устанавливает количественные характеристики информации, определяет физические и статистические параметры каналов связи, формулирует условия согласования источников информации с каналами связи, развивает идеи о применении кодирования для повышения помехоустойчивости передачи по каналам связи с шумом и построении оптимальных кодов для передачи информации в отсутствие шумов и т. д.

Предлагаемое учебное пособие охватывает круг вопросов, соответствующих теоретическим основам передачи информации

в АСУ. Поэтому существенное внимание уделяется как дискретной, так и непрерывной информации, снимаемой с датчиков производственного процесса в АСУ. В общем плане рассматриваются методы выявления детерминированных и случайных сигналов на фоне помех, приводятся рекомендации по выбору оптимальных видов дискретных сигналов и преобразованию непрерывных, способам согласования источников информации с потребителями в сетях связи АСУ. Значительное внимание уделяется информационному подходу к оценке качества функционирования комплекса технических средств АСУ. Анализируется верность преобразования информации в подсистеме сбора и передачи и устанавливаются границы эффективного использования избыточности для обеспечения заданного качества функционирования.

Основу учебного пособия составляют материалы лекций, которые в течение ряда лет читались автором в Ленинградском ордена Ленина электротехническом институте им. В. И. Ульянова (Ленина) для студентов, специализирующихся в области АСУ.

Автор благодарен проф. В. И. Николаеву, а также коллективу кафедры информационно-управляющих систем ЛПИ им. М. И. Калинина, возглавляемому проф. А. И. Яшиным, за ряд ценных замечаний, сделанных при рецензировании рукописи.

# ВВЕДЕНИЕ

## § 1. Автоматизированное управление производством

В современную эпоху научно-технической революции резко возросла роль управления. Управление стало одним из наиболее важных факторов развития производительных сил нашей страны. Уровень организации управления предприятием, отраслью, народным хозяйством в значительной степени определяет эффективность капиталовложений и способствует развитию нашей экономики. Если в последнее время вполне естественной стала автоматизация производства и разработаны основные принципы построения автоматических систем управления, то появление большого объема управленческих работ, усложнение алгоритмов управления предприятием привели к необходимости возникновения автоматизированного управления. Автоматизированная система управления представляет собой совокупность коллектива людей и комплекса технических средств, которые в сочетании с организационным, информационным и математическим обеспечением осуществляют выполнение основной функции управления [8]. Главным при построении АСУ является некоторая цель, реализация которой в процессе создания и эксплуатации системы позволяет получить определенный экономический эффект.

По своему назначению для различных объектов управления к настоящему времени определились следующие группы систем: автоматизированные системы управления организационно-экономического характера, предназначенные для использования при управлении предприятием (АСУП) или отраслью промышленности (ОАСУ), автоматизированные системы управления технологическими процессами (АСУТП), предназначенные для управления конкретным процессом производства продукции и обеспечивающие производство продукции заданного качества, интегральные системы управления, объединяющие внутри себя как АСУП, так и АСУТП. Обобщенная структурная схема

автоматизированной системы управления предприятием представлена на рис. 1.

Автоматизированная система управления обеспечивает управление предприятием на основе осведомляющей информации, которая отображает действительное состояние функционирования предприятия, управляющей информации, вырабатываемой в АСУ с помощью преобразующей информации, включающей в себя нормативные данные, условия преобразования, т. е. определенные алгоритмы и т. д. [11].

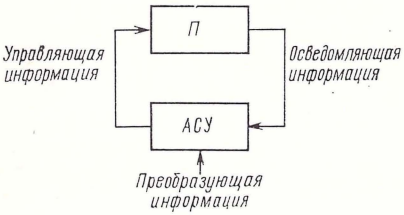


Рис. 1.

Основной информацией для функционирования АСУП является управляющая, которая вырабатывается в автоматизированной системе управления при непосредственном участии человека. Управляющая информация возникает на основе переработки исходных данных по определенным алгоритмам и на основе ряда действий, в настоящее время не поддающихся формализации, и позволяет воздействовать на процесс производства с целью получения заданного качества его функционирования. Основой для получения управляющей информации служат некоторые экономико-математические методы управления, с помощью которых осуществляются обработка информации и принятие оптимальных управляющих решений.

Существенную роль при реализации автоматизированных систем управления играет вычислительная техника. Использование цифровых вычислительных машин в АСУТП позволяет достаточно быстро решать исключительно сложные задачи, связанные с выбором оптимальной стратегии управления, однако наличие средств вычислительной техники не является принципиальной отличительной чертой АСУ.

Автоматизированное управление отличается от автоматического управления той ролью, которую выполняет человек в контуре управления. Если в автоматических системах управления человек выполняет роль диспетчера и действует по детерминированному алгоритму, то в АСУ человек принимает решения. Несмотря на весьма широкое использование средств вычислительной техники, основная роль по выработке оптимального решения остается за человеком. Однако благодаря использованию технических средств человек может достаточно оперативно получить в обобщенном виде интегральную информацию о производстве, обработать ее, дать прогноз производства и на основании этого принять некоторые решения. Для АСУТП характерным является то, что ЦВМ работает в реальном масштабе

времени, в АСУП это имеет место лишь в подсистемах оперативного управления.

Процесс управления в АСУ реализуется с помощью носителей информации. Для АСУП в качестве носителя информации используется документ, в котором информация может представляться в формализованном либо в неформализованном виде. Обычно документ является отражением формализованной информации. Для АСУТП носителем информации является электрический сигнал, который обеспечивает непосредственное воздействие на какие-то исполнительные органы, осуществляющие коррекцию хода технологического процесса с целью обеспечения заданного качества продукции.

ГОСТ 19675—74 определяет АСУ как «человеко-машинную систему, обеспечивающую автоматизированный сбор и обработку информации, необходимой для оптимизации управления в различных сферах человеческой деятельности». Поэтому АСУ могут рассматриваться как некоторые информационные системы, и в зависимости от стадии их разработки и назначения в настоящее время определились следующие ступени развития АСУ:

1) информационно-справочные системы, обеспечивающие выдачу некоторой справочной информации по запросу либо по какому-то иному алгоритму;

2) информационно-советующие системы, осуществляющие подготовку и предварительную обработку определенного объема информации по заданным алгоритмам и тем самым позволяющие улучшить процесс принятия решения;

3) информационно-управляющие системы — системы высшего уровня, в которых отдельные функции управления возлагаются на ЦВМ. Такие системы должны обеспечить наибольший экономический эффект при их внедрении в производство.

Переход к более сложным по выходной информации системам управления вызывает усложнение структуры систем, которые становятся иерархическими как по функциональному принципу, так и по их реализации на технических средствах. Иерархическими становятся структуры комплекса технических средств, на которых строят АСУ. Иерархия системы обусловливает появление достаточно сложных разветвленных сетей связи и сложных информационных потоков, зачастую пересекающихся в отдельных точках сети. Все это требует информационного подхода, знания основ теории информации при создании подобных систем управления производством.

## § 2. Информационный характер процесса принятия решения

В основе функционирования автоматизированной системы управления лежит процесс принятия решения, поскольку АСУ является человеко-машинной системой. Принятие решения о вы-

боре оптимальной стратегии управления и конкретных управляющих воздействий в реально возникшей ситуации возможно на основе некоторого перечня данных, характеризующих как состояние самой управляемой системы, так и состояние среды.

Используемые в настоящее время методы управления и регулирования в АСУ характеризуются достаточной сложностью, что обусловлено невозможностью полного описания системы с помощью какой-то определенной модели, а поэтому всегда возникают такие ситуации, в которых основное решение должен принимать человек. Процесс принятия решения в АСУ имеет ярко выраженный информационный характер. Основные этапы процесса принятия решения отображены на рис. 2. Первым этапом является получение информации (ПИ), характеризующей результат предшествующего управления и состояние внешней среды. Эта информация, получившая название о с в е д о м л я ю щ е й, должна поставляться руководителю предприятия в интегрированном обобщенном виде. Перед получением этой информации должна осуществляться некоторая обработка ее по алгоритмам, которые находят свое отражение в преобразующей информации.

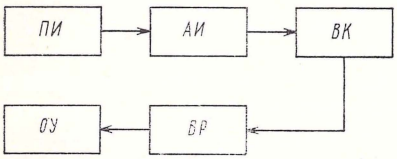

Рис. 2.

Основной целью принятия решения в АСУ в конкретных случаях является ликвидация отклонений в режимах функционирования системы, возникающих под действием различных случайных факторов, связанных с функционированием и самого объекта управления, и системы управления. Они могут быть следствием ошибок оператора, возмущений внешней среды, неисправности технических средств, помех в каналах передачи информации, ошибок в алгоритмах и т. д. Поэтому получение информации, представляющее собой один из основных этапов в информационном процессе принятия решения, является тем исходным фактором, верность которого определяет все остальные этапы.

Полученную информацию необходимо проанализировать. В соответствии с этим появляется этап анализа информации (АИ) с целью выбора альтернативы. При выборе альтернативы существенное место занимает представление всей модели поведения системы. Выбор возможных способов достижения цели в каждом конкретном случае базируется на некоторой модели поведения системы, в связи с чем необходимо выбрать лучший способ достижения цели. Для того чтобы сравнить альтернативы, приходится основываться на определенных критериях, а поэтому весьма существенным является этап, связанный с выбором критерия (ВК) качества функционирования

системы. Должно быть принято такое решение, которое в заданном диапазоне при соответствующих ограничениях обеспечивает требуемое качество функционирования.

На основе критерия и полученной информации о состоянии объекта осуществляется следующий этап — выработка решения (ВР). Следует отметить, что выработка решения является творческим процессом и здесь существенное значение имеют индивидуальные особенности человека, принимающего решения.

Формализация процесса принятия решения возможна только для вполне детерминированных случаев поведения системы, а поскольку за счет внешних возмущений возможно отклонение поведения системы от нормального режима, зачастую возникает необходимость принятия творческого решения, не поддающегося формализации. Управляющее решение для системы организационно-экономического характера выдается в виде документа. В этом документе отображается та управляющая информация, содержание которой по мнению человека, принявшего решение, должно перевести систему в оптимальный режим функционирования.

Следует отметить, что выбор правильного решения и последующая выдача этого решения есть наиболее ответственный этап, при этом весьма важным критерием является не только правильность решения, но и своевременность выдачи его. Поэтому при конкретном управлении производством, особенно в подсистемах АСУ, связанных с оперативным управлением и оперативным регулированием производства, существенное значение имеет быстродействие системы, т. е. своевременность принятого решения. Полученное управляющее решение передается на объект управления (ОУ).

Как видно из рис. 2, процесс принятия решения связан с получением, обработкой, хранением, выдачей информации, и потому подготовка, представление, обработка, хранение, выдача, воспроизведение информации являются теми основными процессами, которые сопутствуют принятию решения, и от быстродействия протекания этих процессов и их верности во многом зависит качество функционирования АСУ в целом.

Оценку качества принимаемых решений в связи с информационным характером процесса принятия решения можно осуществить на основе общих положений теории информации. В соответствии с теорией информации может быть введено понятие функции потерь для каждого дискретного момента времени и определен средний риск при принятии решения. В этом случае принятие решения может быть рассмотрено как процесс, связанный со средним риском, определяемым функцией от характеристик системы. Необходимо при этом найти такую стратегию поведения управляющего устройства, которая обеспечивает минимум среднего риска, т. е. принять решение, при котором средний риск становится минимальным. Тогда основная

постановка задачи и оценка качества принятия решения в значительной степени аналогичны задачам теории связи, учитывающим наличие помех в исходной информации, в некотором канале и в устройстве выдачи информации. Поскольку принятие решения осуществляется на основе подготовленной информации, а эта информация циркулирует в конкретной АСУ на базе комплекса технических средств (КТС), то представляется целесообразным рассмотреть обобщенно состав КТС.

## § 3. Основные подсистемы комплекса технических средств АСУ

Функционирование автоматизированной системы управления базируется на совокупности технических средств, связанных между собой единством передаваемой информации и информационных ограничений, что составляет комплекс технических средств АСУ. При построении этого комплекса принимается ряд критериев, которые являются глобальными критериями оценки выбранного варианта построения и связаны с приведенными затратами по созданию конкретного варианта КТС, а также ряд частных критериев, определяющих собой технические характеристики системы, такие, как верность переработки информации, время ее переработки, надежность технических средств и т. д.

В настоящее время промышленностью разработаны различные виды технических средств, которые можно объединять путем комплексирования на основе единой технической базы. В этом комплексе могут быть выделены средства подготовки информации, обеспечивающие автоматический учет и регистрацию информации на различных носителях, в том числе на машинных, средства сбора информации по различным типам каналов связи с разными характеристиками входных потоков, средства передачи информации, которые разделяются по типам используемых каналов связи и составляют в основном различные варианты модемов (совокупности модулятора и демодулятора) и устройств защиты от ошибок, средства обмена информацией, представляющие собой оконечные установки-концентраторы информации, осуществляющие обмен информацией с электронными вычислительными машинами, средства выдачи и воспроизведения информации, к которым относятся различного вида экраны, индикаторные табло и т. д.

Все перечисленные виды технических средств можно объединить в единый комплекс и представить в виде отдельных подсистем, связанных между собой линиями передачи информации (рис. 3). В общем случае различают следующие основные подсистемы КТС: регистрации и подготовки информации (ПРПИ), сбора и передачи информации (ПСПИ), хранения

и обработки информации (ПХОИ), выдачи и воспроизведения информации (ПВВИ) [17].

Как видно из рис. 3, непосредственно с объектом управления (ОУ) связаны подсистема регистрации и подготовки информации и подсистема выдачи и воспроизведения информации. С объекта управления информация поступает в ПРПИ, где принимает вид осведомляющей информации и в зависимости от расположения информационно-вычислительного центра либо передается непосредственно на информационно-вычислительный центр, который составляет ПХОИ, либо поступает в ПСПИ при достаточной удаленности объектов управления от центра управления. В соответствии с этим различают подсистемы сбора и передачи с курьерской и с автоматической связями. Учитывая дальнейшее развитие АСУ, мы рассматриваем здесь структуру подсистемы с автоматической связью, которая осуществляется по специально выделенным или по уплотненным линиям связи.

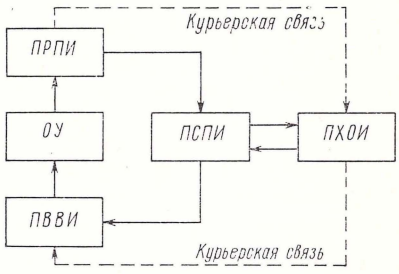


Рис. 3.

Информация, поступающая в ПХОИ, в зависимости от алгоритмов преобразования, которые закладываются в программах ЭВМ, может храниться определенное время и обрабатываться в соответствии с поступающей извне преобразующей информацией. В данной подсистеме элементом, принимающим решение, может быть человек, что приводит к образованию некоторого человеко-машинного комплекса.

На основе предварительно обрабатываемой информации и решений, принимаемых или по детерминированным алгоритмам или с помощью человека, формируется управляющая информация, которая курьером либо посредством автоматической связи ПСПИ выдается на ПВВИ. Эта подсистема оказывает управляющие воздействия на объект управления, в этом качестве может выступать предприятие, и тогда АСУ имеет организационно-экономический характер, а носителем информации в основных подсистемах выступает документ. Выходная информация фиксируется в документе. Если объектом управления является технологический процесс, то получаем КТС, который функционирует в структуре АСУТП. Тогда выходной информацией является электрический сигнал, воздействующий на технологический процесс таким образом, чтобы получилось заданное качество выпускаемой продукции. Вместе с тем воспроизведение информации дает возможность оператору, который находится на ОУ или в центре управления, осуществлять

оперативный контроль за качеством функционирования систем. Посредством устройств воспроизведения информации может быть получена информация о ходе производства, а также различные данные, выдаваемые с определенным интервалом времени в форме документов.

Особенность структуры КТС АСУ при построении его на базе стандартных технических средств заключается в необходимости увязки технических средств, поэтому в основу комплекса должна быть положена одна и та же конструктивно-технологическая база, которая сопрягалась бы с конкретным типом используемой вычислительной машины. В организационно-экономических системах нашли применение цифровые вычислительные машины. Для АСУТП характерным является использование управляющих вычислительных машин, которые в совокупности с цифроаналоговыми преобразователями позволяют осуществить автоматизацию процесса управления.

Общая компоновка комплекса технических средств в зависимости от назначения системы может осуществляться по-разному. Например, некоторые технические средства, имеющиеся в настоящее время, совмещают функции регистрации, сбора и передачи информации, другие могут объединять в себе функции хранения, обработки и выдачи информации, но всегда в структуре комплекса можно выделить основные подсистемы, что позволяет методически грамотно подойти как к выбору отдельных составляющих КТС, так и к их сопряжению.

В настоящее время при построении КТС основное внимание уделяется качеству функционирования, которое оценивается двумя критериями — временем преобразования информации в каждой подсистеме и верностью преобразования ее. Нередко при оценке комплекса могут быть использованы информационные характеристики самого объекта. В целом КТС представляет собой техническое обеспечение, которое совместно с математическим, информационным и организационным обеспечением составляет совокупность обеспечивающих подсистем АСУ и позволяет при правильном проектировании получить эффект от каждой функциональной подсистемы АСУ.

## § 4. Место и роль теории информации при разработке АСУ

Теория информации базируется на разработанной К. Шенноном «Математической теории связи» [28]. Развитые в этой работе идеи кодирования при помехоустойчивой передаче сообщений по каналам с шумами, вопросы сокращения избыточности, понятия каналов с шумом стали основополагающими в современной теории информации. Сформулированным К. Шенноном фундаментальным теоремам в дальнейшем было дано строгое математическое доказательство и на основе их разви-

тия получены весьма значительные результаты в области алгебраической теории кодирования, построения различных оптимальных кодов, в области декодирования избыточных кодов, получения границ вероятности ошибки и различных информационных пределов избыточности. Вместе с тем теория информации стала применяться не только при построении систем передачи информации. Ее идеи проникли в сферу хранения информации, где существенное место занимали задачи минимизации представления информации, минимизации аппаратуры, используемой для хранения и обработки информации.

Методы, разработанные в теории информации, начинают использоваться и для оценки качества функционирования системы. В частности, при анализе случайных воздействий на передаваемую информацию модель канала с шумом переносится на функционирование технических средств, где канала передачи информации как такового не существует, но вводится условный канал с шумом отказов аппаратуры системы, что позволяет получить ряд характерных закономерностей, присущих функционированию технических систем [4].

Таким образом, в современном виде теория информации представляет собой научную дисциплину, которая изучает способы передачи и хранения информации в наиболее надежном и удобном виде. При построении АСУ теория информации позволяет грамотно строить процессы, связанные с преобразованием информации в комплексе технических средств, а именно наиболее экономно представить и подготовить информацию, осуществить контроль над информацией при минимальной ее избыточности, а также достоверный ввод информации, передачу по каналам связи и хранение с минимальными потерями, обработать информацию с минимальными ошибками арифметическими способами контроля и, наконец, воспроизвести информацию в наиболее удобном и оптимальном виде. Иначе говоря, можно осуществить выбор оптимального варианта с учетом достижений современной теории информации. Весьма существенным оказывается также выбор единицы измерения информации и количественных характеристик, которые позволяют сравнить полезную и вредную шумовую информацию, определить потери информации в каждой подсистеме и найти методы оптимизации построения отдельных подсистем КТС с целью снижения потерь информации при ее преобразовании.

В теории передачи информации до сих пор сохраняют актуальность вопросы, связанные с оптимальным использованием каналов связи, во-первых, с экономической эффективностью использования их, а во-вторых — с наиболее полным использованием физических характеристик канала, согласования канала связи с сигналом и получения оптимальных вариантов его преобразования, т. е. оптимальных вариантов кодирования, оптимальных методов модуляции и т. д. Поскольку в АСУ исполь-

зуются сложные информационные сети обмена информацией, то возникают принципиальные задачи распределения информационного потока по сети таким образом, чтобы качество функционирования ее находилось на заданном уровне [14]. Для этого могут быть использованы основные положения теории информации, определяющие соотношения между скоростью передачи информации и пропускной способностью канала связи, и могут быть найдены методы, позволяющие повысить скорость передачи при сохранении требований по верности.

В АСУ широко применяется как дискретная, так и непрерывная информация. Проблемы представления дискретной информации возникают в каналах с шумом при решении задачи выбора необходимой избыточности, которая должна быть увязана с требуемыми вероятностно-временными характеристиками функционирования системы. Эта задача реализуется в ПСПИ и может быть решена только при полном использовании основ теории информации и их развитии в конкретном прикладном направлении.

Весьма существенными являются проблемы кодирования информации при подготовке и регистрации, где информация должна быть представлена в наиболее экономном виде и с достаточной степенью точности. В этом случае можно использовать положения по оптимальному кодированию информации в каналах связи без шума с учетом статистических свойств источника. На этапе преобразования непрерывной информации возникают проблемы дискретизации ее и выбора оптимальной шкалы квантования по времени и по уровню. При передаче информации существенным является соотношение скорости передачи и пропускной способности непрерывного канала. При любом преобразовании информации в каждой подсистеме присутствуют помехи, они могут быть весьма различными в зависимости от используемых технических средств, от разветвленности информационной сети [27]. Поэтому возникают задачи оптимального выделения дискретной и непрерывной информации на фоне помех, поскольку в этих случаях могут быть использованы совершенно различные методы их отсеивания.

Применение теории информации для оценки качества функционирования систем привело в настоящее время к проблеме комплексной оценки качества с учетом как процесса передачи информации, помех в каналах связи, так и процесса функционирования технических средств АСУ, где возмущающими воздействиями выступают неисправности аппаратуры, приводящие к различным последствиям: постепенным или мгновенным отказам или сбоям [15].

Присутствие человека в структуре АСУ приводит к отдельным ошибкам, что ухудшает качество функционирования системы. Возникает задача описания человека-оператора, при этом могут быть использованы результаты, полученные в теории

информации, или созданы информационные модели человека-оператора, что позволяет оценить качество функционирования системы.

Не меньшее значение результаты теории информации имеют и в области информационного обеспечения АСУ, определения оптимальной структуры информационных массивов. Минимизация длин отдельных блоков представляемой информации в массиве является одной из актуальных задач. Верность хранения информации, организация информационного описания объектов управления, построение информационно-логических моделей АСУ — все это базируется на современной теории информации.

Таким образом, совершенно ясно, что в современном понимании теория информации является исключительно емкой дисциплиной, позволяющей технически грамотно подходить к основам построения автоматизированных систем управления с точки зрения как функциональной структуры АСУ, так и оптимальной структуры технического и информационного обеспечения.

*Глава 1*

# ИНФОРМАЦИОННЫЙ ПРОЦЕСС В АСУ

## § 1. Фазы преобразования информации в АСУ

Основой функционирования автоматизированной системы управления является информационный процесс, характеризующийся определенными фазами преобразования информации, основные из которых нашли отражение в подсистемах комплекса технических средств.

Среди фаз обращения информации следует отметить прежде всего фазу подготовки информации. Подготовка информации может осуществляться вручную, машинным способом, с использованием различных видов носителей. Для увеличения быстродействия целесообразно осуществлять этап подготовки, используя либо машинные носители информации, либо такие сигналы, которые способны передаваться непосредственно от источника в канал связи. На этапе подготовки информация, снимаемая с объекта управления или подготовленная в результате действий оператора, наносится на некоторый носитель по определенному правилу и далее включается в информационный процесс. Другой фазой является регистрация информации, осуществляемая с целью образования документа, в котором информация дана в формализованном виде. Этот документ может храниться и использоваться при последующем управлении.

Следующими фазами преобразования являются сбор и передача информации. Сбор информации обычно осуществляется с территориально разнесенных точек объекта управления, например в пределах цеха, предприятия, отрасли промышленности, а также и с более отдаленных объектов. Сбор информации нередко включает в себя и неразрывно связан с передачей информации, которая состоит в переносе информации на значительные расстояния посредством дополнительного преобразования исходных сообществ в сигналы, способные по своим физическим свойствам передаваться по выбран-

ным каналам связи. При передаче информации формируются дополнительные сигналы, а поэтому на приемной стороне должны использоваться специальные средства, осуществляющие выявление сигналов по определенным алгоритмам.

Информация, передаваемая по каналам, в дальнейшем используется при принятии решения, поэтому она должна быть обработана. Обработка информации в АСУ производится с помощью ЭЦВМ, где централизуются функции обработки, и на основе отдельных моделей ситуации человеком с помощью детерминированных или неформализованных способов осуществляется принятие решения. При этом информация подвергается цифровым и аналоговым преобразованиям. В процессе обработки возможны промежуточные этапы хранения информации с использованием оперативных и долговременных запоминающих устройств, построенных на различных технических средствах. На этапе хранения информации возникает весьма серьезная задача систематизации имеющейся информации. Из информации формируется набор данных, создается банк данных. В итоге возникают проблемы создания информационных массивов, а также поиска и организации информационных массивов таким образом, чтобы обеспечить заданное быстродействие при записи и выводе информации из массива.

Поскольку выработка управляющих воздействий немыслима без дальнейшего использования информации в контуре управления, то важной задачей является вывод информации в соответствующем виде и ее воспроизведение. Воспроизведение информации требуется прежде всего тогда, когда в информационном процессе активное участие принимает человек и ему необходимо получить качественные и количественные характеристики выходной информации в наглядной форме. Для этого используются специальные технические средства цифрового и графического характера, а также экраны с различного типа мнемосхемами, индикаторами и т. д., которые способны воздействовать на органы чувств человека. Кроме того, информация может быть выдана в виде документов, которые далее используются в АСУ организационно-экономического характера. Отдельные данные из управляющей вычислительной машины с помощью цифроаналоговых преобразователей могут выдаваться непосредственно в виде электрических сигналов, которые осуществляют либо регулирующие, либо управляющие, либо защитные действия и тем самым корректируют технологический процесс.

В качестве объекта управления может выступать как производство, так и отдельный технологический процесс. При этом, с одной стороны, объект управления является источником информации, а с другой — ее потребителем, т. е. информационная система является замкнутой, в которой имеют место внутренние и внешние возмущающие воздействия.

868805

Наряду с крупными этапами или фазами преобразования информации существуют более мелкие операции, связанные с отдельными воздействиями на информацию для получения каких-то данных по заранее известным алгоритмам. Сюда прежде всего можно отнести такую дополнительную операцию, как классификация. Классификация оказывается необходимой в ряде случаев, например, при хранении информации, когда данные, накапливаемые случайным образом, должны храниться в форме, удобной для последующего их извлечения. При этом выбираются определенные классификационные признаки, которые вносятся в саму информацию и хранятся вместе с основной информацией. Весьма существенной операцией является синтез. Данная операция необходима в случае, когда требуется объединить отдельные составляющие данные по одному и тому же вопросу в совокупность данных для получения единого логически связанного слова. Независимо от фазы преобразования информации каждый вид ее обладает определенными характеристиками, среди которых полезно выделить связанные с функционированием АСУ следующие характеристики.

1. *Цель информации.* Под целью информации можно понимать назначение процесса информирования, выработку и применение решения, выдачу команд и т. д.

2. *Способ передачи и формат информации.* Следует отметить, что формат информации бывает различным в смысле приемлемости по отношению к человеку и машине. Целесообразно при этом, чтобы человек получал большую часть информации в виде документа, где записан и систематизирован материал. В настоящее время все большее применение находят средства отображения — дисплеи, и они в некоторой степени определяют тот формат информации, который может быть выдан человеку-оператору.

3. *Избыточность информации.* При подготовке информации возникает определенная избыточность, связанная с периодичностью ввода информации, особенно при вводе отдельных точек непрерывной величины. Избыточность может иметь и источник дискретной информации, когда отдельные выдаваемые сообщения взаимно зависимы. Наличие исходной избыточности уменьшает быстродействие системы, увеличивает форматы сообщений, и поэтому одна из задач, которую приходится решать,— это устранение первичной избыточности в информации. Вместе с тем в фазе передачи информации избыточность является средством, полезным для борьбы с внешними возмущающими воздействиями (помехами), и при правильном выборе уровня вводимой избыточности и алгоритма построения сигнала в канале связи получаем повышение достоверности передачи информации по каналу.

4. *Время преобразования информации* является одной из основных характеристик функционирования комплекса технических средств. Оно зависит как от формата информации, который определяется свойствами технических средств, так и от алгоритма управления информационными потоками в информационной сети АСУ. В целом возникают временные задержки, которые вызывают старение информации и снижение ее ценности, если информация используется в системе оперативного управления производством. В связи с этим быстродействие есть одна из важнейших характеристик подсистем комплекса технических средств АСУ.

5. *Периодичность появления информации* зависит непосредственно от конкретной функциональной подсистемы, по которой вводится информация. Периодичность может изменяться, при этом существуют информация, которая должна обрабатываться в реальном масштабе времени, а также информация, которая может обрабатываться спустя значительный отрезок времени после ее возникновения, что особенно характерно для подсистем перспективного планирования, технической подготовки производства и др.

6. *Верность информации* — одна из основных характеристик информации на любой фазе ее преобразования. В зависимости от верности информации определяется степень доверия к информации в процессе принятия решения, и верностью исходных данных определяется эффективность функционирования АСУ. Поэтому подготовка информации, точно отражающей конкретный производственный процесс, является одним из актуальных вопросов при проектировании самой системы. Вместе с тем построение комплекса технических средств АСУ, а также математического и информационного обеспечения в значительной степени осуществляется с учетом требований верности информации при функционировании АСУ. Таким образом, наличие многих фаз преобразования информации и отдельных операций по преобразованию на каждой фазе вызывает появление различных форм представления информации в машине и для человека-оператора, но во всех случаях информация должна поставляться верно и своевременно.

## § 2. Понятие сообщения и кода

Сведения о состоянии объекта управления в АСУ формируются в виде сообщений. Под сообщением понимается все то, что подлежит передаче. Независимо от содержания сообщение обычно представляется в виде электрического, звукового, светового, механического или других сигналов [23]. Таким образом, сообщение отображает некоторые исходные сигналы любого вида и по свойствам зависит от исходных сигналов.

В АСУ все исходные сигналы, поступающие от объекта, можно разделить на две большие группы: сигналы статические, которые отображают устойчивые состояния некоторых объектов и могут быть представлены, например, в виде определенного положения элемента системы, текста в документе, определенного состояния электронного устройства и т. д., и сигналы динамические, для которых характерно быстрое изменение во времени, отображающее, например, изменение электрических параметров системы.

Динамические и статические сигналы имеют свои области использования. Статические сигналы существенное место занимают при подготовке, регистрации и хранении информации. Динамические используются в основном для передачи информации. Однако заметим, что это не всегда является обязательным.

По характеру изменения сигналов во времени различают сигналы непрерывные и дискретные. Непрерывный сигнал отображается некоторой непрерывной функцией и физически представляет собой непрерывно изменяющиеся значения колебаний. Дискретный сигнал характеризуется конечным множеством значений и в зависимости от исходного состояния принимает значения, связанные с определенным состоянием системы. Исходя из физической сущности процесса, свойственного объекту управления, можно выделить некоторые разновидности непрерывных и дискретных функций, отображающих реальные сигналы [26]:

1) непрерывную функцию непрерывного аргумента. Функция имеет вид $f(t)$, непрерывна на всем отрезке и может описать реальный сигнал в любой момент времени. При этом не накладывается никаких ограничений на выбор момента времени и на выбор значения самой функции;

2) непрерывную функцию дискретного аргумента. Обычно такие сигналы возникают при квантовании непрерывных величин по времени. В этом случае задаются некоторые фиксированные моменты времени $t_i$, отсчитываемые через интервал $\Delta t$, который обычно определяется спектральными свойствами исходного физического процесса. Функция $f(t_i)$ может принимать любые мгновенные значения, но она определяется лишь для дискретных значений времени. Этот вид сигналов и связанных с ним функций имеет место при формировании исходных сообщений из непрерывных величин;

3) дискретную функцию непрерывного аргумента $f_j(t)$. В этом случае функция имеет ряд конечных дискретных значений, однако определена на всем отрезке времени $t$ для любого мгновенного значения времени. Дискретизация самой функции связана с созданием шкалы квантования по уровню, что свойственно различным датчикам, при этом шаг

квантования определяется требуемой точностью воспроизведения исходной величины;

4) дискретную функцию дискретного аргумента $f_j(t_i)$. В этом случае функция принимает одно из возможных дискретных значений, общее число которых является конечным, и определяется для конечного набора дискретных значений времени. Имеем дискретизацию как по уровням, так и по моментам времени.

В целях систематизации сообщений и обеспечения возможности передачи сообщений по каналам связи используются процедуры кодирования. С помощью кодирования сообщение представляется в форме, которая позволяет осуществить передачу его по каналам связи. Дискретное сообщение можно изобразить в виде некоторой последовательности цифр или букв, при этом каждая цифра или буква представляет собой одно сообщение. С помощью кода каждая цифра или буква отображаются некоторым набором импульсов, которые составляют кодовую комбинацию. Основное требование, предъявляемое к кодовым комбинациям, состоит в возможности различения их на приемной стороне при определенных воздействиях помех в каналах связи. Общее число кодовых комбинаций равно числу возможных сообщений $M$.

При построении кода учитывается ряд особенностей, связанных с возможностями передачи информации по каналу связи, кроме того, вопрос реализации технических средств преобразования сообщений в код, т. е. построение кодирующих устройств и соответствующих им средств обратного преобразования — декодирующих устройств. Весьма важными являются вопросы обеспечения требуемой верности и скорости передачи информации [29]. В настоящее время в различных системах передачи информации и в том числе в информационных сетях АСУ получило распространение большое число кодов, а поэтому рассмотрим их обобщенную классификацию.

1. *По основанию системы счисления* коды делятся на двоичные, троичные, четверичные и т. д. В каждой системе счисления используется определенная совокупность символов, при этом число возможных символов для К-ичной системы счисления равно $K$. Двоичные коды строятся с помощью символов 0,1; троичные — 0,1, 2, при этом нуль означает отсутствие передачи информации по каналу, т. е. отсутствие импульса, единица означает символ с одним значением сигнального признака, двойка — с другим. Под сигнальным признаком понимается некоторое значение тока или напряжения, позволяющее отличить один символ от другого.

2. *По построению* коды делятся на систематические и несистематические. Особенность построения систематических кодов как разделимых заключается в том, что в них

четко разделены часть кода, несущая основную информацию, и часть кода, служащая для обнаружения и исправления ошибок, которая представляет собой контрольную информацию. Систематические коды могут быть построены по детерминированным алгоритмам, в соответствии с чем можно осуществить достаточно простые способы выявления этих кодов с обнаружением или исправлением ошибок.

Несистематические коды строятся с использованием различных методов комбинирования. Это коды на одно сочетание, коды на размещения, на перестановки и т. д., и при их выявлении осуществляется анализ кода путем сопоставления принятой комбинации с известным набором кодов на приемной стороне.

3. *По наличию избыточности* коды делятся на избыточные и неизбыточные. Для неизбыточных кодов характерно то, что при каждом отображении сообщения кодовой комбинацией для числа $M$ возможных кодовых комбинаций основным свойством является возможность их различения. Тогда код при основании системы счисления $K$ может быть построен как отображение множества десятичных чисел от нуля до $M-1$ с числом разрядов $n$ в каждой кодовой комбинации. Например, для $M=4$ двоичный неизбыточный код может быть получен как представление чисел 0, 1, 2, 3 двухэлементным двоичным кодом: 00,01,10,11 соответственно. Переход от $K$-ичного числа к десятичному можно осуществить по формуле

$$N=\sum_{i=0}^{n-1}\alpha_i K^i,$$

где $n$ — число элементов в коде, или длина кода; $K$ — основание системы счисления кода; $\alpha_i$ — значение символа в $i$-м разряде, причем младшим является разряд, расположенный справа. Следует отметить, что символы кода в линии связи передаются в обратном порядке, т. е. сначала старший разряд и далее остальные.

Если необходимо представить, например, четыре сообщения троичным неизбыточным кодом, то исходные десятичные числа 0,1,2,3 запишутся в виде 00,01,02,10. В общем случае $m$-элементным неизбыточным кодом в $K$-ичной системе счисления можно представить $M=K^m$ сообщений. Например, при двухэлементном неизбыточном троичном коде можно иметь $3^2=9$ сообщений.

Переход от неизбыточного кода к избыточному при использовании систематических кодов осуществляется путем добавления некоторых контрольных позиций, которые можно получить либо путем различных логических операций, выполняемых над основными информационными позициями, либо путем использования детерминированных алгоритмов, связы-

вающих неизбыточный и избыточный коды. Например, если нужно перейти от неизбыточного кода к простейшему избыточному, то для случая двоичного кода, рассчитанного на четыре сообщения, отображением которых являются кодовые комбинации 00,01,10,11, достаточно ввести одну контрольную позицию, значение символа на которой будет определяться как сумма значений предшествующих символов по модулю два. Эта логическая операция в двоичной системе определяется равенствами $0 \oplus 0 = 0$, $1 \oplus 1 = 0$, $0 \oplus 1 = 1$, $1 \oplus 0 = 1$. Для рассматриваемых сообщений получаем 000, 011, 101, 110. Особенность такого кода заключается в том, что он позволяет обнаружить любую одиночную ошибку. Таким образом, отличие неизбыточных кодов от избыточных состоит в том, что из-за отсутствия избыточности они не способны обнаруживать ошибки и поэтому не могут быть использованы для передачи информации по каналам с шумом. С целью обеспечения достоверной передачи информации по каналу связи при заданных вероятностно-временных ограничениях необходимо вводить избыточность в код, что можно осуществить путем использования дополнительных контрольных позиций.

4. *По корректирующим свойствам* коды делятся на обнаруживающие и исправляющие, или корректирующие. Обнаруживающие коды при введении в них избыточности позволяют обнаруживать ошибки, с помощью корректирующих кодов возможно исправление ошибок, при этом доля вводимой избыточности по сравнению с предыдущим существенно возрастает. Следует отметить, что в настоящее время в АСУ при передаче информации более широко применяются обнаруживающие коды в сочетании с дополнительными алгоритмами повышения помехоустойчивости за счет использования обратного канала связи.

5. *По расположению элементов* кода во времени различают последовательные, параллельные и последовательно-параллельные коды. В АСУ чаще применяются коды с последовательной передачей элементов во времени в связи с особенностями использования средств модуляции и демодуляции в каналах связи. Трудность реализации параллельных кодов заключается в том, что должны быть использованы либо такие сигнальные признаки (например, частотный), которые допускают одновременную передачу нескольких своих значений, либо совокупность сигнальных признаков при одновременной передаче по одному значению каждого сигнального признака.

В качестве примера рассмотрим возможность параллельной передачи нескольких сообщений при использовании амплитудного, полярного и сигнального признаков по длительности. Тогда если требуется передать три сообщения, одно из которых передается импульсом большой амплитуды, другое

импульсом отрицательной полярности, а третье импульсом большой длительности, то передача этих сообщений будет означать появление импульса большой амплитуды, отрицательной полярности и большой длительности. Очевидно, нетрудно рассмотреть возможность передачи любой совокупности двух сообщений одновременно.

Следует отметить, что параллельные коды могут быть эффективно использованы при передаче относительно небольших объемов информации, так как основанные на них модемы обладают малой скоростью передачи.

## § 3. Кодирование информации в АСУ

Процесс кодирования информации осуществляется практически во всех подсистемах комплекса технических средств АСУ, однако в зависимости от фаз преобразования информации используются коды различных типов и разные способы представления информации.

Рассмотрим последовательно основные подсистемы комплекса технических средств и процесс кодирования в них. Для подсистемы подготовки и регистрации информации (ПРПИ) характерным является преобразование исходного сообщения, которое имеет дискретный или непрерывный характер, в неизбыточный код. При этом в зависимости от статистики сообщения используются коды одинаковой либо неодинаковой длины. Если все комбинации кода имеют одну и ту же длину, то коды называются равномерными, и в этом случае любое сообщение отображается одним и тем же числом символов в кодовой комбинации. При неравновероятных исходных сообщениях возникает проблема отображения их с помощью кодов неодинаковой длины, получивших название неравномерных.

В зависимости от характера исходных сообщений могут быть использованы коды, различные по своему построению [20]. Если реализуется процесс преобразования непрерывной величины в код, то применяют специфичные коды, которые обеспечивают наибольшую точность преобразования. Примером такого кода может считаться код Грея, относящийся к числу двоичных кодов. Он используется обычно для отображения десятичных чисел, поэтому применяется шестнадцать возможных комбинаций его, представляющих собой четырехзначные двоичные последовательности. Основная особенность кода Грея заключается в том, что при аналого-кодовом преобразовании возникает ошибка считывания. Она будет наименьшей, если каждая последующая кодовая комбинация отличается от предыдущей в одном разряде, что позволяет свести ошибку к единице в младшем разряде.

Десятичные числа и соответствующий им код Грея представлены ниже.

| Десятичное число | Код Грея | Десятичное число | Код Грея | Десятичное число | Код Грея | Десятичное число | Код Грея |
|---|---|---|---|---|---|---|---|
| 0 | 0000 | 4 | 0110 | 8 | 1100 | 12 | 1010 |
| 1 | 0001 | 5 | 0111 | 9 | 1101 | 13 | 1011 |
| 2 | 0011 | 6 | 0101 | 10 | 1111 | 14 | 1001 |
| 3 | 0010 | 7 | 0100 | 11 | 1110 | 15 | 1000 |

Каждая последующая кодовая комбинация отличается от предыдущей лишь в одном разряде. В соответствии с этим уменьшается ошибка при считывании, и код Грея, таким образом, позволяет реализовать аналого-кодовое преобразование при минимальной ошибке считывания.

Наряду с кодом Грея при подготовке информации применяются и другие коды. Прежде всего это двоично-десятичные коды, употребляющиеся в случае, когда необходимо представить многоразрядные десятичные числа в двоичной системе счисления. Тогда каждая цифра десятичного числа записывается в виде четырехразрядного двоичного числа. Наибольшее применение получил двоично-десятичный код с весами чисел 8, 4, 2, 1, т. е. $K^i$, где $K = 2$. Кодовые комбинации для этого случая представлены ниже.

| Десятичное число | Двоично-десятичный код | Десятичное число | Двоично-десятичный код | Десятичное число | Двоично-десятичный код |
|---|---|---|---|---|---|
| 0 | 0000 0000 | 4 | 0000 0100 | 8 | 0000 1000 |
| 1 | 0000 0001 | 5 | 0000 0101 | 9 | 0000 1001 |
| 2 | 0000 0010 | 6 | 0000 0110 | 10 | 0001 0000 |
| 3 | 0000 0011 | 7 | 0000 0111 | | |

Обычно код используют в качестве промежуточного при введении информации в вычислительную машину, если исходное сообщение представлено десятичным кодом. Перевод чисел из десятичной формы в двоично-десятичную может осуществляться автоматически, для этого можно использовать перфоратор.

Значительную группу составляют коды, используемые для подготовки информации к вводу в ЦВМ. В машинах первого поколения при вводе буквенно-цифрового текста использовался пятиразрядный международный код, который по мере

развития цифровой вычислительной техники стал расширяться, появились варианты шести-, семиразрядных кодов. В 1964 г. был введен ГОСТ на стандартный семиразрядный код. В настоящее время принят новый семиразрядный международный первичный код.

Таким образом, особенностью ПРПИ является то, что в ней возможно применение неизбыточных кодов, которые при равновероятности входных сообщений имеют постоянную длину. Если входные сообщения неравновероятны, то возникает проблема уменьшения средней длины кода за счет согласования источника информации со свойствами полученных кодовых комбинаций. В этом случае основная идея кодирования сводится к тому, чтобы наиболее часто встречающимся сообщениям присвоить наиболее короткие кодовые комбинации. Тогда уменьшается средняя длина кода и увеличивается количество информации, которая переносится в одном элементе кода.

Для ПСПИ характерно использование кодов с избыточностью. В этом случае основная характеристика кодов — это минимальное кодовое расстояние. Кодовое расстояние между соседними комбинациями определяется числом различающихся разрядов. Если имеется совокупность кодовых комбинаций, отображающих $M$ сообщений, то кодовое расстояние определяется по наименьшему кодовому расстоянию, которое имеет место среди возможных комбинаций. Обычно кодовое расстояние геометрически интерпретируется с помощью графа кода [30].

На рис. 4 в качестве примера представлен граф простейшего двухэлементного неизбыточного кода, для которого число сообщений $M = K^n = 2^2 = 4$, а кодовые комбинации имеют вид 00, 01, 10, 11. Каждая комбинация отображается *узлом* (вершиной) графа. Узлы графа соединяют *ребра*. Минимальное число ребер графа между заданными кодовыми комбинациями называется *числом переходов* и обозначается $d$. Если $d = 1$, то код является неизбыточным, и любое воздействие помехи на него может привести к трансформации сообщения, т. е. код не обладает помехоустойчивостью. Введем в код избыточность, т. е. выберем из числа сообщений $M = 4$ две комбинации, наиболее удаленные друг от друга. В этом случае код содержит две рабочие и две запрещенные комбинации. Приемная схема реагирует лишь на рабочие комбинации. Тогда $M = 2$ и $d = 2$, код — избыточный, и при воздействии одной ошибки возникает защитный отказ, т. е. обеспечивается обнаружение одиночной ошибки. При наличии двух ошибок в данном коде происходит трансформация сообщения. В общем случае для кодов, обнаруживающих ошибки, $d = r + 1$, где $r$ — число обнаруживаемых ошибок.

Рассмотрим двоичный трехэлементный код. Комбинации неизбыточного кода имеют следующий вид: 000, 001, 010, 011, 100, 101, 110, 111. Граф представлен на рис. 5. Исследуем свойства этого кода. Если используются все комбинации, то $M = 8$, $d = 1$, $r = 0$. Для повышения помехоустойчивости выбираем половину кодовых комбинаций с четным или нечетным числом единиц, тогда $M = 4$, $d = 2$, $r = 1$.

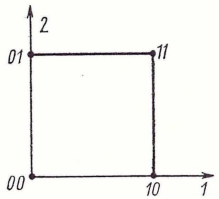


Рис. 4.

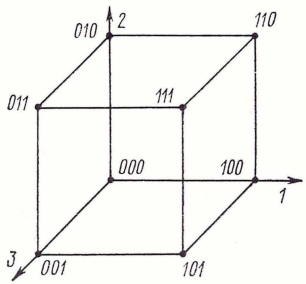


Рис. 5.

Перейдем к коду, исправляющему ошибки. Для этого из всего множества комбинаций выберем две, наиболее удаленные друг от друга, например, 001 и 110, т. е. $M = 2$, $d = 3$, $r = 2$. Допустим, что приемная схема рассчитана на прием только этих комбинаций, тогда возникает обнаруживающий код, причем возможно обнаружение двух ошибок.

Рассмотрим, как происходит исправление ошибки. Пусть тот же код используется для передачи двух сообщений. Будем считать, что разрешенными являются рабочие комбинации, а также все другие, отстоящие от них на один переход. Реализуем приемную схему следующим образом: введем параллельные цепи в декодирующее устройство так, чтобы комбинации 001, 101, 011, 000 вызывали срабатывание первого исполнительного элемента 1И, а комбинации 110, 010, 100, 111 — второго исполнительного элемента 2И. Пусть передается первое сообщение кодом 001 и возникает одиночная ошибка вида 100, в результате которой получается новая комбинация 101. Тогда срабатывает исполнительное устройство 1И, и ошибка исправляется. При любой одиночной ошибке происходит отступление на один переход от рабочей комбинации, т. е. ошибка исправляется. При двух ошибках происходит трансформация сообщения. Например, если передается комбинация 001 и возникают две ошибки вида 110, то срабатывает исполнительный элемент 2И. В этом случае получаем код $M = 2$, $d = 3$, $r = 1$, $s = 1$, где $s$ — число исправляемых ошибок. Число кодовых переходов $d = r + s + 1$, причем $r \geqslant s$.

Сравним по помехоустойчивости обнаруживающий и исправляющий коды, построенные с одинаковым кодовым расстоянием $d$. В качестве критериев помехоустойчивости примем вероятность прохождения сообщения $P_{пр}$, вероятность трансформации (преобразования) сообщения под действием помех $P_{тр}$, вероятность защитного отказа $P_{зо}$, которая охватывает события, связанные с пропаданием сообщения в канале связи. Очевидно, что $P_{пр} + P_{тр} + P_{зо} = 1$, причем вероятность ошибки $P_{о} = 1 - P_{пр} = P_{тр} + P_{зо}$.

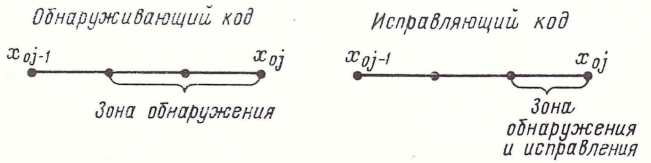


Рис. 6.

При переходе от обнаруживающего кода к исправляющему увеличивается вероятность прохождения сообщения, но снижается его обнаруживающая способность, и поэтому увеличивается вероятность трансформации. Следовательно, при переходе к корректирующему коду вероятность ошибки снижается, но возрастает доля вероятности трансформации. На рис. 6 показаны соотношения между обнаруживающим и исправляющим кодами для комбинаций $x_{oj-1}$, $x_{oj}$ при кодовом расстоянии $d = 3$.

Чтобы эффективно использовать корректирующий код, необходимо компенсировать увеличение вероятности трансформации либо за счет увеличения обнаруживающих свойств кода, либо за счет накопления их на приемной стороне [16]. Практическое применение в ПСПИ нашли в настоящее время коды циклического типа, исправляющие ошибки. Для простоты схемной реализации нередко используются коды, обнаруживающие ошибки, с повторением в системе с обратной связью. Если возможность повторной передачи отсутствует, то могут быть использованы достаточно просто реализуемые коды, исправляющие ошибки малой кратности.

В ПХОИ наиболее часто применяются коды для арифметических операций. В этом случае используются следующие основные правила выполнения арифметических действий в двоичной системе счисления. Для сложения: $0+0=0$, $0+1=1$, $1+0=1$, $1+1=10$; для вычитания: $0-0=0$, $1-0=1$, $1-1=0$, $10-1=1$: для умножения: $0\times 0=0$, $0\times 1=0$, $1\times 0=0$, $1\times 1=1$. По этим правилам можно сложить числа, отображаемые в двоичной системе счисления, и полу-

чить ответ, тождественный десятичной системе. Например, складывая 3 и 4, получаем в десятичной системе $3+4=7$, в двоичной системе $011+100=111$, что в десятичной системе означает цифру 7. Таким образом, используя основные правила арифметических действий в двоичной системе счисления, мы можем осуществить обработку информации.

Весьма важной задачей при реализации кодов в ПХОИ является контроль над выполнением арифметических действий. В этом случае могут применяться различные виды контроля, в том числе введение определенной избыточности в коды, с помощью чего удается осуществить контроль правильности выполнения отдельных операций.

При хранении информации используются коды, позволяющие осуществить систематизацию и классификацию информации, в частности неизбыточные коды, отображающие или отдельные сообщения, или их совокупности, которые позволяют в определенном порядке (обычно в порядке естественного возрастания чисел) располагать информационные записи в массиве и тем самым упрощают в дальнейшем их поиск и выведение информации из массива.

Для ПВВИ характерно представление информации в такой форме, которая доступна оператору и может быть использована для последующего управления. Поэтому в случаях, когда необходимо представить информацию оператору, целесообразно применять десятичную систему счисления и осуществлять преобразование выходной информации подсистемы сбора и передачи в десятичный код, для чего можно использовать специальные декодирующие устройства по выявлению каждого сообщения. В соответствии с этим для выявления буквенного текста необходимо использовать на приемной стороне ПСПИ устройство декодирования, позволяющее по принятой совокупности элементов с любой избыточностью и любым основанием системы счисления осуществить однозначное выявление текста в понятной для человека форме. На практике это осуществляется техническими средствами с помощью матриц выявления кодовых комбинаций. При воспроизведении информации на табло необходимо применять специальные индикаторы, которые в настоящее время строятся в десятичной системе счисления, тогда получим информацию в форме, доступной для оперативного использования. Регистрация информации при ее выдаче из ЦВМ осуществляется также в обычном цифро-буквенном тексте, и поэтому применение каких-либо дополнительных кодов оказывается нецелесообразным.

На всех этапах кодирования, на всех фазах преобразования информации в АСУ важное значение имеет вопрос верности, и в этом смысле кодирование является одним из эффективных средств повышения верности преобразования информации как непосредственно в самом процессе преобразования, так и при

вводе, подготовке и выявлении информации. Использование избыточных кодов позволяет существенно повысить верность информации, снижающуюся за счет внутренних и внешних возмущающих воздействий в системе. Эту избыточность в определенной степени можно использовать как избыточность, вводимую в структуру системы. Тогда удается повысить надежность функционирования КТС, что также способствует обеспечению заданной верности выходной информации в АСУ.

## § 4. Информационные сети и информационные потоки

В АСУ большое место при функционировании КТС занимают процессы сбора и обработки информации, а также выдачи управляющего воздействия. Поэтому возникает автономная сеть обмена информацией, которая является внешней по отношению к источникам и потребителям информации, т. е. возникает некоторая информационная сеть, осуществляющая обмен информацией с заданными вероятностно-временными характеристиками доведения информации до потребителя, обслуживание информационных потоков в точках объединения или пересечения информации, а также обеспечивающая верность передачи информации по каналам связи.

По отношению к информационной сети объекты управления представляют собой абоненты, при этом внешними оказываются источники и потребители информации, которые подсоединяются к входам и выходам информационной сети [10]. Таким образом, по отношению к информационной сети информация, поступающая от источников, является внешним управляющим воздействием, но наряду с ними в информационной сети возникают внешние возмущающие воздействия, которые проявляются в виде отказов, сбоев аппаратуры сети, помех, действующих в каналах связи. Эти воздействия имеют случайный характер и могут описываться некоторыми вероятностными закономерностями.

Во временном аспекте информация, поступающая в сеть, представляет собой информационный поток, характеризующийся моментами появления отдельных сообщений, и нередко этот поток является случайным и характеризуется статистическими закономерностями. Для оценки характеристик сети можно использовать характеристики, свойственные любой системе передачи информации, которые увязывают между собой верность и время передачи. В широком смысле верность передачи информации включает в себя верность, связанную с непосредственной передачей информации по каналам связи, и вероятностную оценку исправности технических средств, на которых строится информационная сеть. Таким образом, информационная сеть базируется на комплексе технических

средств, а поэтому он должен удовлетворять соответствующим вероятностно-временным характеристикам.

Формализуя информационную сеть, можно выделить в ней узлы, где осуществляются ввод, выдача или коммутация информации. Узлы информационной сети связаны дугами, отображающими собой каналы связи. Получаем граф, в вершинах которого располагаются источники, потребители или узлы коммутации информации, а дуги представляют собой линии или каналы связи. Поскольку информационная сеть должна удовлетворять определенным условиям быстродействия, то весьма существенным фактором, определяющим качество функционирования информационной сети АСУ, является метод коммутации, который применяется для переключения и объединения информационных потоков в отдельных промежуточных узлах сети. В настоящее время по методу коммутации все сети разделяют на три вида: 1) с коммутацией каналов; 2) с коммутацией сообщений; 3) с комбинированной коммутацией.

Методом коммутации каналов реализуется предварительное установление электрического соединения между источником и потребителем информации, после чего осуществляется использование данной линии связи лишь для передачи этой информации. При передаче информации от данного источника другие источники, желающие передать информацию по данной линии связи, должны ожидать освобождения канала, т. е. момента, когда коммутация будет осуществлена на их линию связи. Использование метода коммутации каналов предполагает наличие отказов в обслуживании со стороны информационной сети отдельным источникам. В задачах АСУ отказ означает задержку информации во времени, что может привести к потере ценности информации, так как старение информации практически означает потерю информации для управления. Более оптимальное использование информационной сети наблюдается в случае, когда линия связи не закрепляется за определенным источником информации и источник может осуществлять выдачу информации по мере ее появления до какого-то следующего узла сети, и далее по мере возможности эта информация передается последовательно от одного коммутационного узла к другому путем коммутации сообщений.

Метод коммутации сообщений осуществляется без организации предварительного соединения, и в этом случае приходится каждому сообщению присваивать адрес, в зависимости от которого сообщение проходит набор коммутационных пунктов. При занятости каналов связи в коммутационных точках осуществляется хранение информации, сообщение ставится в очередь и по мере освобождения канала связи передается по следующим каналам.

Метод коммутации сообщений является исключительно удобным для передачи информации в АСУ. В связи с тем, что

на разных уровнях используются каналы связи с различными физическими и статистическими характеристиками, использование коммутационных узлов позволяет осуществить трансформацию скоростей передачи информации в соответствии с используемыми типами каналов связи, видами модуляции и устройствами защиты от ошибок. В промежуточных узлах (коммутационных центрах) удается осуществить компоновку и хранение сообщений, поступающих по низкоскоростным каналам связи от источников информации, и далее выдать скомпонованные блоки сообщений по высокоскоростным каналам, т. е. осуществить оптимальное использование высокоскоростных каналов связи. Метод коммутации сообщений позволяет также осуществить приоритетное обслуживание отдельных сообщений, имеющих особую важность при управлении производством, в соответствии с чем удается осуществить задержку менее приоритетных сообщений и выполнить первоочередное обслуживание сообщений, связанных с управляющей информацией.

Нетрудно убедиться в том, что реализация метода коммутации сообщений требует дополнительного объема оборудования, усложняет алгоритм управления информационными потоками в сети, на что, однако, приходится идти в связи с той эффективностью, которую обеспечивает этот метод в информационных сетях АСУ. В отдельных случаях, когда необходимо осуществить и передачу сообщений, отображающих различные виды информации, и передачу разговорного текста, возможно совмещение метода коммутации каналов и метода коммутации сообщений. В результате получаем информационную сеть с комбинированным способом коммутации.

Поскольку метод коммутации сообщений является наиболее приемлемым для сети в АСУ, рассмотрим более подробно информационную сеть с коммутацией сообщений. Модель информационной сети с коммутацией сообщений представлена на рис. 7. Сеть отображается некоторым графом, в котором в оконечных узлах располагаются источники информации (ИИ) и потребители информации (ПИ), а в промежуточных узлах располагаются коммутационные центры, в данном случае узлы коммутации сообщений (УКС). Информация от ИИ поступает на УКС. Поступление информации от ИИ осуществляется по мере появления ее и обычно не связано с управляющими воздействиями от УКС.

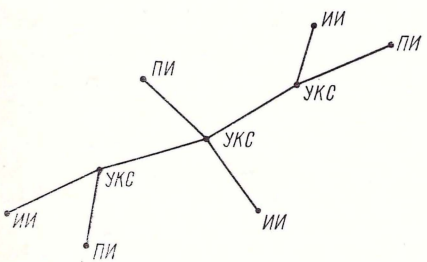


Рис. 7.

В ряде случаев можно предусмотреть случай циклического опроса источников информации от УКС. Тогда возникает проблема уменьшения времени опроса, определения числа источников в одной опрашиваемой группе с целью оптимизации вероятностно-временных характеристик обслуживания. Сообщения в УКС хранятся в виде блоков, и по определенному алгоритму, который зависит от структуры сети и закладывается при ее проектировании, передаются по высокоскоростному каналу связи, соединяющему УКС. Количество промежуточных УКС определяется структурой АСУ, разветвленностью ее основных организационных подразделений, а поэтому в ряде случаев информация может проходить несколько узлов коммутации до поступления на потребитель. От последнего УКС информация по отдельной линии связи передается на потребитель (источник). Обычно для этого используются низкоскоростные каналы связи, выполненные в виде специально выделенных либо уплотненных линий.

Для обеспечения заданного качества функционирования системы приходится оптимизировать информационную сеть по ряду критериев. Обобщенным критерием оптимизации информационной сети является критерий приведенных затрат, который учитывает единовременные капитальные вложения, а также затраты, распределенные во времени и связанные с эксплуатацией данного оборудования, с обслуживающим персоналом и т. д. Однако требование удешевления информационной сети должно быть разумным, оно должно сочетаться с требованиями по частным критериям, характеризующим качество функционирования технических средств сети, куда относятся верность передачи информации по сети и время передачи. Верность передачи информации в информационной сети определяется возможностями обслуживания информационных потоков в узлах коммутации. Здесь возникает вероятность отказа в обслуживании, вероятность потери информации. Верность передачи обусловлена также действиями помех в каналах связи. Если на нижнем уровне специально выделенных каналов при расчете верности могут быть использованы модели независимых ошибок, пуассоновских потоков помех, то на уровне высокоскоростных каналов связи в настоящее время используются более сложные модели ошибок, учитывающие группирование ошибок в пачки, возможность перекрытия отдельных пачек ошибок, где в определенной степени можно лишь условно использовать экспоненциальный закон для интервалов между пачками ошибок. Начиная с модели Гильберта, широкое применение нашли модели ошибок, связанные с несколькими состояниями канала связи.

На верность обработки и выявления информации существенное влияние может оказать неисправность технических средств, следствием которой может быть появление отказов и сбоев в аппаратуре. Одной из характеристик функционирования явля-

ется время обслуживания. Оно складывается из времени передачи информации в канале связи, времени обработки информации в ЦВМ, а также времени задержки. Последнее связано с организацией передачи информации по методу коммутации сообщений, когда при группировке сообщений из-за занятости верхнего канала могут возникать значительные задержки, имеющие в целом случайный характер и оценивающиеся статистическими закономерностями. Время задержки зависит от дисциплины обслуживания сообщений в узлах коммутации, от организации передачи служебной информации, адресации информации и т. д.

Таким образом, информационная сеть АСУ представляет собой широко разветвленную сеть, в которой могут возникнуть различные случайные возмущающие воздействия и которая по отношению к потребителям информации является внешней сетью, обладающей своими показателями качества функционирования. Топология информационной сети определяется структурой используемых линий связи. По структуре сети можно разделить на радиальные, кустовые, цепочечные и смешанные. Варианты структур линий связи представлены на рис. 8. Источники информации в сети создают во времени информационные потоки. В зависимости от типа и структуры технических средств источника информации, располагаемого в оконечном пункте сети, возможны различные модели информационных потоков во времени. Наиболее часто встречающейся моделью потока является регулярный поток, характеризующийся частотой появления сообщений во времени. Более общей моделью потока является случайный поток. Простейшей моделью случайного потока может служить пуассоновская модель, из которой нетрудно установить вероятность возникновения заданного числа сообщений на фиксированном интервале времени. Однако пуассоновский поток не всегда может быть использован как модель источника информации, поэтому при практической реализации и проектировании сети иногда необходимо знать реальную структуру потока, для чего проводят специальные экспериментальные исследования источников и оценивают их статистические свойства. Расчет характеристик сети можно осуществить аналитическим способом или дополнить моделированием, но результаты расчета без экспериментального обследования источников информации будут приближенными и лишь в определенной степени соответствовать реальной сети.

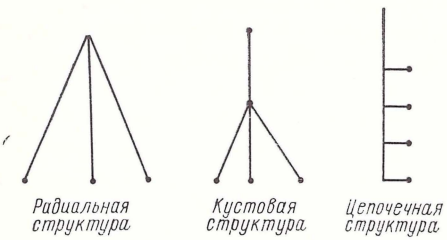


Рис. 8.

## § 5. Виды модуляции и их реализация в информационных сетях АСУ

Для информационных сетей, используемых при обмене информации в АСУ, характерно существование различных скоростей передачи информации на отдельных участках, связывающих источники информации, коммутационные центры и потребители информации. Высокоскоростными каналами являются каналы связи между центрами коммутации сообщений, где должны быть использованы такие средства преобразования сигнала, которые позволяют наилучшим образом использовать свойства канала связи.

Основным средством преобразования сигнала, которое позволяет передавать сигнал с необходимой скоростью и с минимальным затуханием по каналу связи, является модуляция [25]. Модуляцией называется воздействие на некоторый параметр переносчика с целью передачи управляющего сигнала по данному каналу. В настоящее время применяются непрерывные и дискретные переносчики. Непрерывные переносчики бывают двух видов: постоянные и синусоидальные. В качестве дискретного переносчика обычно используется некоторая последовательность импульсов.

Наибольшее практическое применение нашел непрерывный переносчик. При модуляции его импульсами кода для постоянного переносчика можно получить либо простой позиционный, либо полярный позиционный код. Эту процедуру называют манипуляцией постоянного переносчика, и спектр получаемой последовательности практически мало отличается от спектра исходного управляющего импульса. Значительное повышение помехоустойчивости может быть получено, если использовать непрерывный переносчик, с помощью которого осуществляется сдвиг несущей частоты в сторону частот канала связи, которые пропускаются с наименьшими затуханиями.

Воздействуя на непрерывный переносчик, можно получить следующие основные виды модуляции. При воздействии на амплитуду гармонического колебания получаем амплитудную модуляцию (АМ), на частоту — частотную модуляцию (ЧМ), на фазу — фазовую модуляцию (ФМ). Особенностью использования различных видов модуляции является изменение ширины спектра исходного управляющего сигнала, представляющего собой некоторую последовательность импульсов, отображающих код одного или нескольких сообщений. В соответствии с этим представляется целесообразным оценить спектр сигналов, передаваемых с помощью различных типов модуляции.

1. В качестве переносчика при амплитудной модуляции используется колебание, уравнение которого можно записать в виде $u = U\cos(\omega_0 t + \varphi_0)$. Поскольку начальный угол фаз $\varphi_0$

не имеет принципиального значения, положим его в дальнейшем равным нулю. Амплитуда переносчика меняется по закону $U = U_0(1 + m_{AM}\, x(t))$, где $m_{AM} = \Delta U/U_0$ — коэффициент амплитудной модуляции, $x(t)$ — модулирующее воздействие. Тогда если представить простейший случай модуляции в виде некоторого косинусоидального колебания, для которого $x(t) = \cos \omega_1 t$, то напряжение, модулированное по амплитуде: $U_{AM} = U_0 (1 + m_{AM} \cos \omega_1 t) \cos \omega_0 t$, или

$$u_{AM} = U_0 \cos \omega_0 t + (m_{AM}/2)\, U_0 \cos (\omega_0 - \omega_1)\, t + \\ + (m_{AM}/2)\, U_0 \cos (\omega_0 + \omega_1)\, t.$$

На рис. 9 представлено модулирующее воздействие (*а*), амплитудно-модулированное напряжение (*б*) и спектр (*в*,*г*). Видно, что полученный спектр (*г*) включает в себя три гармонические составляющие с амплитудами: $U_0$, представляющей собой основную несущую частоту переносчика $\omega_0$, и $m_{AM}U_0/2$ — амплитудами боковых частот. Спектр АМ колебания при гармоническом управляющем воздействии является линейчатым. Важным вопросом при использовании АМ модуляции является выбор коэффициента $m_{AM}$. Наиболее глубокая модуляция получается при $m_{AM} = 1$, и в этом случае мощность АМ колебания в полтора раза превышает мощность переносчика. Обычно выбирается $m_{AM} = 1/2$.

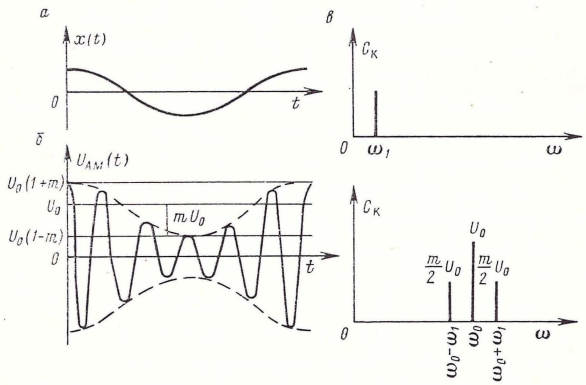


Рис. 9.

В общем случае управляющее воздействие представляет собой достаточно сложную модулирующую функцию и может быть представлено в виде ряда Фурье. Тогда $x(t) = \sum_{k=1}^{\infty} C_k \cos (k\omega_1 t - \varphi_k)$, и, подставляя $x(t)$ в формулу для

напряжения амплитудно-модулированного колебания, получаем

$$u_{\text{АМ}} = U_0 \left[ 1 + m_{\text{АМ}} \sum_{k=1}^{\infty} C_k \cos (k\omega_1 t - \varphi_k) \right] \cos \omega_0 t =$$

$$= U_0 \left\{ \cos \omega_0 t + \frac{m_{\text{АМ}}}{2} \sum_{k=1}^{\infty} C_k \cos [(\omega_0 + k\omega_1) t - \varphi_k] + \right.$$

$$\left. + \frac{m_{\text{АМ}}}{2} \sum_{k=1}^{\infty} C_k \cos [(\omega_0 - k\omega_1) t + \varphi_k] \right\}.$$

Поскольку спектр управляющего воздействия является сложным линейчатым спектром, состоящим из ряда гармонических составляющих, то и спектр амплитудно-модулированного колебания получается также сложным и включает колебания с частотами $\omega_0$, $\omega_0 + k\omega_1$ и $\omega_0 - k\omega_1$, т. е. получаем основное гармоническое колебание с частотой $\omega_0$ и две боковые полосы, включающие в себя колебания с суммарными и разностными частотами. Использование амплитудной модуляции увеличивает ширину спектра в два раза, что обеспечивает повышение помехоустойчивости (см. рис. 9). С другой стороны, за счет модуляции реализуется сдвиг на величину $\omega_0$ спектра исходного сигнала и осуществляется возможность передачи модулированного сигнала по каналу связи в области полосы пропускания с наименьшим затуханием.

2. Для частотной модуляции характерным является воздействие модулирующего колебания на частоту переносчика. В этом случае частота переносчика меняется по закону $\omega = \omega_0 + \Delta\omega x(t)$, где $x(t)$ — модулирующее воздействие; $\Delta\omega$ — наибольшее отклонение частоты при модуляции, это отклонение называют *девиацией* частоты.

Рассмотрим простейший случай, когда модулирующее воздействие представляет собой гармонический сигнал, для которого $x(t) = \cos \omega_1 t$. Напряжение частотно-модулированного колебания можно записать в виде $u_{\text{ЧМ}} = U \cos \varphi$. Учитывая, что $U = U_0$, $\varphi = \int \omega dt + \varphi_0$, и полагая $\varphi_0 = 0$, получаем

$$\begin{aligned} u_{\text{ЧМ}} &= U_0 \cos \left\{ \int [\omega_0 + \Delta\omega x(t)]\, dt \right\} = \\ &= U_0 \cos \left[ \omega_0 t + \Delta\omega \int x(t)\, dt \right] = \\ &= U_0 \cos \left[ \omega_0 t + \Delta\omega \int \cos \omega_1 t dt \right] = \\ &= U_0 \cos \left[ \omega_0 t + (\Delta\omega/\omega_1) \sin \omega_1 t \right] = \\ &= U_0 \cos (\omega_0 t + \beta \sin \omega_1 t), \end{aligned}$$

где $\beta$ — индекс частотной модуляции. Обычно $\beta \gg 1$.

Учитывая возможность разложения тригонометрических функций в ряд с помощью функций Бесселя, при котором

$$\cos(\beta \sin \omega_1 t) = I_0(\beta) + 2\sum_{k=1}^{\infty} I_{2k}(\beta) \cos 2k\omega_1 t,$$
$$\sin(\beta \sin \omega_1 t) = 2\sum_{k=1}^{\infty} I_{2k+1}(\beta) \sin(2k+1)\omega_1 t,$$

где $I_k(\beta)$ — функция Бесселя порядка $k$ от аргумента $\beta$, определяемая по таблицам, получаем

$$u_{\text{чм}} = U_0 \{I_0(\beta) \cos \omega_0 t + \sum_{k=1}^{\infty} I_k(\beta) \cos(\omega_0 + k\beta) t + \\ + \sum_{k=1}^{\infty} (-1)^k I_k(\beta) \cos(\omega_0 - k\beta) t\}.$$

Амплитудный спектр напряжения при частотной модуляции зависит от значения $\beta$ и теоретически является бесконечным. Оценим граничные значения $\beta$. Если $\beta \to 0$, то $\cos(\beta \sin \omega_1 t) \to 1$, $\sin(\beta \sin \omega_1 t) \to \beta \sin \omega_1 t$, и напряжение частотно-модулированного колебания может быть записано в виде

$$u_{\text{чм}} = U_0 \cos \omega_0 t - U_0 \beta \sin \omega_0 t \sin \omega_1 t = \\ = U_0 \cos \omega_0 t + (\beta/2) U_0 \cos(\omega_0 + \omega_1) t - (\beta/2) U_0 \cos(\omega_0 - \omega_1) t.$$

В этом случае ширина спектра $\Delta\omega_c = 2\omega_1$, т. е. получаем, что при $\beta \to 0$ частотная модуляция по своим свойствам и занимаемой полосе спектра практически не отличается от амплитудной модуляции.

Если $\beta > 1$, то необходимо знать значения $I_k(\beta)$, которые можно найти из графиков изменения функции Бесселя (рис. 10).

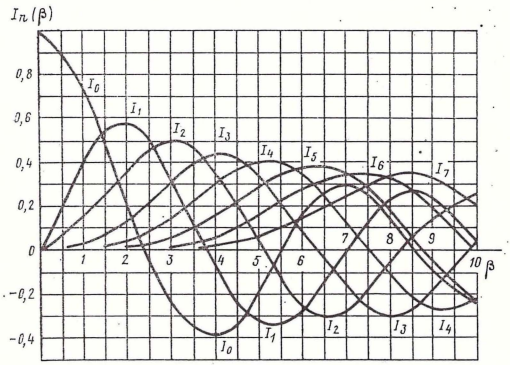


Рис. 10.

Видно, что с увеличением $\beta$ значения $I_k(\beta)$ убывают. При $k > \beta$ эти значения становятся очень малыми, а поэтому амплитудами этих составляющих спектра колебаний можно пренебречь. Если ограничить число возможных боковых частот

величиной $n = \beta + 1$, то ширина спектра может быть определена как

$$\Delta\omega_c = 2n\omega_1 = 2(\beta + 1)\omega_1 = 2(\Delta\omega + \omega_1).$$

Если $\beta \gg 1$, то $\Delta\omega_c \approx 2\Delta\omega$, и ширина спектра в основном определяется лишь девиацией частоты, т. е. индексом частотной модуляции, и практически не зависит от частоты модулирующего воздействия.

При рассмотрении модулирующего воздействия негармонического типа получим аналогичные соотношения по изменению спектра при переходе к частотной модуляции, а поэтому можно лишь отметить основные особенности частотной модуляции. Частотная модуляция при $\beta \gg 1$ позволяет реализовать спектр с шириной полосы в $\beta$ раз большей, чем спектр сигнала при амплитудной модуляции, что позволяет обеспечить большую помехоустойчивость передачи. Использование частотной модуляции, однако, требует широкополосных каналов, и при частотном уплотнении каналов имеем меньше каналов связи в одной и той же линии, чем при амплитудной модуляции. Частотная модуляция при своей реализации требует более сложной и дорогостоящей аппаратуры, чем амплитудная.

3. Особенностью фазовой модуляции является то, что управляющее воздействие изменяет фазу переносчика. В этом случае фаза сигнала изменяется по закону $\varphi = \omega_0 t + \Delta\varphi x(t) - \varphi_0$, где $\Delta\varphi$ представляет собой девиацию фазы и зависит от амплитуды управляющего воздействия. Рассмотрим простейший случай управляющего воздействия, представляющего собой гармоническое колебание. Пусть имеет место управляющее воздействие $x(t) = \cos\omega_1 t$. Для простоты будем считать $\varphi_0 = 0$, тогда

$$u_{\text{ФМ}} = U_0 \cos\varphi = U_0 \cos[\omega_0 t + \Delta\varphi x(t)] =$$

$$= U_0 \cos[\omega_0 t + \Delta\varphi \sin(\omega_1 t + \pi/2)].$$

Сравним фазовую модуляцию с частотной, для этого определим значение частоты сигнала в виде $\omega = d\varphi/dt = \omega_0 - \Delta\varphi\omega_1 \sin\omega_1 t$.

Если $\Delta\varphi\omega_1 = \Delta\omega$ и $\Delta\omega$ — девиация частоты частотной модуляции, то нетрудно установить пределы изменения частоты гармонических составляющих спектра сигнала. В спектре частотно-модулированного сигнала появляются отдельные гармонические составляющие с частотами от $\omega_{\min} = \omega_0 - \Delta\varphi\omega_1$ до $\omega_{\max} = \omega_0 + \Delta\varphi\omega_1$. Спектр напряжения фазомодулированного колебания совпадает в этом случае со спектром напряжения частотно-модулированного колебания, а ширина спектра $\Delta\omega_c = 2\Delta\varphi\omega_1$. Особенностью спектральных составляющих фазомодулированного колебания является то, что ширина спектра

зависит от частоты модулирующего сигнала. При изменении модулирующей частоты число спектральных линий в спектре остается постоянным, интервалы между этими линиями изменяются, а поэтому изменяется и общая ширина спектра.

Если сравнить фазовую модуляцию с предшествующими видами модуляций, то можно указать, что по своим свойствам она обладает большей помехоустойчивостью, однако требует более сложной схемной реализации выявителей фазы на приемной стороне. В АСУ при создании модемов применяются различные виды фазовой модуляции, а также совместные способы воздействия на параметры переносчика.

При использовании дискретного переносчика воздействию могут подвергаться амплитуда, частота, фаза и длительность импульсов переносчика. В соответствии с этим возникают амплитудно-импульсная (АИМ), частотно-импульсная (ЧИМ), фазо-импульсная (ФИМ) и широтно-импульсная (ШИМ) модуляции. При импульсной модуляции переносчик представляет собой последовательность импульсов и характеризуется линейчатым спектром, вокруг каждой спектральной линии переносчика возникают боковые частоты. Протяженность боковых частот по ширине зависит от вида используемой модуляции, кроме того, может меняться и амплитуда отдельных частотных составляющих в полосе боковых частот. Для обеспечения малого затухания при передаче импульсных последовательностей по каналу связи нередко сами импульсы представляют собой управляющее воздействие для некоторого непрерывного переносчика, и в итоге получаем, что модулированная последовательность модулирует некоторое высокочастотное несущее колебание.

В основе определения спектра различных видов импульсной модуляции лежит знание спектра немодулированной последовательности импульсов. При определении его необходимо учитывать, что немодулированная последовательность $u(t)$ имеет вид

$$u(t) = C_0 + \sum_{k=1}^{\infty} C_k \cos k\omega_0 t = \frac{U\tau}{T}\left(1 + \sum_{k=1}^{\infty} \frac{2\sin k\omega_0\tau/2}{k\omega_0\tau/2} \times \right.$$

$$\left. \times \cos k\omega_0 t\right) = \frac{U\tau}{T}\left[1 + \sum_{k=1}^{\infty} d_k \cos k\omega_0 t\right],$$

где $U$, $\tau$, $T$ — амплитуда, длительность и период импульсов переносчика, $d_k$ — амплитуда $k$-й гармонической составляющей.

Для оценки спектра различного вида импульсной модуляции необходимо использовать ранее приведенную методику.

Рассмотрим в качестве примера спектр амплитудно-импульсной модуляции. Если положить, что модулирующая функция имеет вид $\cos \omega_1 t$, то амплитуда импульсов меняется по закону $u = U(1 + m_{\text{АМ}} \cos \omega_1 t)$. Тогда

$$z(t) = \frac{U\tau}{T}(1 + m_{\text{АМ}} \cos \omega_1 t)\left[1 + \sum_{k=1}^{\infty} d_k \cos k\omega_0 t\right] =$$

$$= \frac{U\tau}{T}\left[1 + m_{\text{АМ}} \cos \omega_1 t + \sum_{k=1}^{\infty} d_k \cos k\omega_0 t + \right.$$

$$\left. + \sum_{k=1}^{\infty} m_{\text{АМ}} d_k \cos \omega_1 t \cos k\omega_0 t\right] = \frac{U\tau}{T}\left[1 + \right.$$

$$+ m_{\text{АМ}} \cos \omega_1 t + \sum_{k=1}^{\infty} d_k \cos k\omega_0 t +$$

$$\left. + \sum_{k=1}^{\infty} \frac{m_{\text{АМ}} d_k}{2} \cos (k\omega_0 + \omega_1) t + \sum_{k=1}^{\infty} \frac{m_{\text{АМ}} d_k}{2} \cos (k\omega_0 - \omega_1) t\right].$$

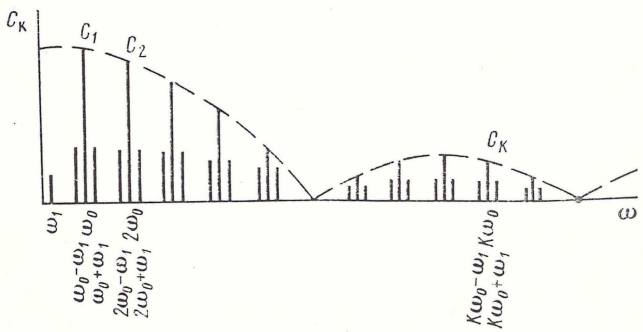


Рис. 11.

Спектр полученного амплитудно-импульсного колебания представлен на рис. 11. Видно, что в спектре появляются составляющие с частотами $k\omega_0 \pm \omega_1$, отражающие наличие модулирующего воздействия. Если модулирующее воздействие имеет более сложный характер и не соответствует гармоническому колебанию, то вблизи каждой составляющей спектра немодулированных импульсов могут возникать боковые полосы, повторяющие спектр модулирующего воздействия. Аналогичным образом можно найти спектры ЧИМ, ФИМ, ШИМ.

В связи с тем, что импульсная модуляция в настоящее время не нашла широкого применения для передачи в информационных сетях АСУ, более подробно на этом вопросе останавливаться не будем. Рассмотрим различные виды модуляции при конкретной реализации системы передачи информации в АСУ. Модуляция реализуется в системах на основе технических средств, представляющих собой стандартные средства в виде модемов. В настоящее время в модемах чаще используют непрерывный переносчик, реализующийся на базе постоянного или синусоидального сигнала. В соответствии с этим построены модемы, в которых применяются двухпозиционные полярные сигналы, если реализуется постоянный переносчик. При реализации переменного переносчика синусоидального типа строятся модемы с использованием амплитудной, частотной или фазовой модуляции [10].

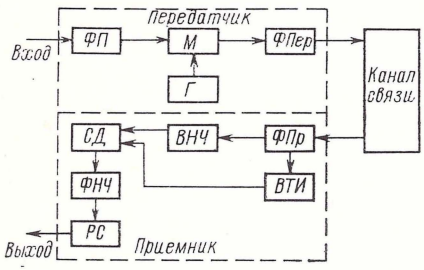


Рис. 12.

Рассмотрим обобщенную структурную схему модема, не зависящую от типа модуляции (рис. 12). Модем состоит из передатчика и приемника. Основными частями передатчика являются формирователь посылок (ФП), осуществляющий формирование посылок заданной формы, модулятор (М), осуществляющий модуляцию заданного типа, фильтр передатчика (ФПер), обеспечивающий пропускание сигнала заданного спектра, и генератор (Г). Таким образом, с помощью этих основных узлов происходит формирование модулированного колебания, которое поступает далее в канал связи. В приемнике основными элементами модема являются следующие: фильтр приемника (ФПр), настроенный на полосу частот спектра передаваемого сигнала, выявитель несущей частоты, или выявитель частоты переносчика (ВНЧ), синхронный детектор (СД), фильтр низкой частоты (ФНЧ), выявляющий огибающую сигнала, выявитель тактовых импульсов (ВТИ), осуществляющий выявление элементов синхронизации, и решающая схема (РС), обеспечивающая различение возможных значений сигнального признака. Использование стандартных устройств типа модемов позволяет обеспечить высокоскоростную передачу информации по выбранным каналам связи на значительное расстояние в информационных сетях АСУ.

## *Глава II*

# КОЛИЧЕСТВЕННЫЕ ХАРАКТЕРИСТИКИ ИНФОРМАЦИИ

## § 1. Количество информации в равновероятных сообщениях

При оценке количества информации прежде всего возникает вопрос о виде исходной информации, а поэтому измерение информации в значительной степени зависит от подхода к самому понятию информации, т. е. от подхода к ее содержанию. В настоящее время существуют три основные теории, в которых к понятию содержательного характера информации подходят с разных позиций [20]. Статистическая теория оценивает информацию с точки зрения меры неопределенности, снимаемой при получении информации. Как правило, она не затрагивает смысла передаваемой информации, т. е. ее семантического содержания. В статистической теории основное внимание обращается на распределение вероятностей отдельных квантов информации и построение на его основе некоторых обобщенных характеристик, позволяющих оценить количество информации в каком-то кванте.

Совершенно иной подход наблюдается в семантической теории, которая учитывает в основном ценность информации, полезность ее и тем самым помогает связать ценность информации со старением, ценность информации и количество ее — с эффективностью управления в системе. Наконец, структурная теория рассматривает структуру построения отдельных информационных массивов, при этом за единицу информации принимаются некоторые элементарные структурные единицы — кванты, и количество информации оценивается простейшим подсчетом квантов в информационном массиве.

Выбор единицы информации в настоящее время является весьма актуальной задачей. При передаче непрерывных сообщений зачастую используется их дискретизация во времени, поэтому применяется геометрическая мера, позволяющая определить количество информации в отдельных отсчетах, снимаемых за некоторый интервал времени, т. е. количество передаваемых сообщений в этом случае определяется числом отсчетов.

При передаче дискретной информации простейшей мерой информации может служить число кодовых комбинаций, отображающих передаваемые сообщения. Число комбинаций получается на основе комбинаторного метода и определяется структурой построения кода, его избыточностью, т. е. способом построения. Недостатком данной меры является нелинейная зависимость между числом кодовых комбинаций и числом элементов в коде. Например, для неизбыточного кода число кодовых комбинаций $M = K^n$. Обычно по каналу связи передается последовательность $n$ символов, поэтому целесообразно иметь характеристику, линейно связанную с числом элементов в коде.

Будем считать, что число сведений $f$ в сообщении линейно зависит от длины кода: $f = kn$. Формулу для числа сведений в сообщении выведем при следующих условиях: 1) осуществляется передача дискретных сообщений; 2) сообщения являются равновероятными и взаимно независимыми; 3) символы, выдаваемые источником, взаимно независимы; 4) система счисления конечна [12]. Тогда $df = kn$. Если $M = K^n$, то $dM = K^n \ln K dn$, $dn = dM/K^n \ln K$ и $df = k dM/M \ln K$;

$$f = k \ln M/\ln K = k_1 \log_a M/\ln K = k_0 \log_a M,$$

где $k_0 = k_1/\ln K$.

*В теории информации за единицу количества информации принято число сведений, которое передается двумя равновероятными символами или сообщениями. Эта единица называется* двоичной единицей информации.

Учитывая сказанное, имеем при $f = 1$ и $M = 2$: $1 = k_0 \log_a 2$. Если $k_0 = 1$, то $a = 2$, $f = I = \log_2 M$, где $I$ — количество информации в некотором усредненном сообщении. Формула $I = \log_2 M$ называется формулой Хартли, она справедлива в соответствии с принятыми выше ограничениями 1) — 4) (см. с. 20).

Рассмотрим, как влияет на число сведений основание кода. Пусть $M$ сообщений передается двумя кодами с основаниями $K_1$ и $K_2$ и числами элементов $n_1$ и $n_2$. Будем считать, что оба кода передают одинаковое число сведений, т. е. $M = K_1^{n_1} = K_2^{n_2}$, тогда $k(K_1)n_1 = k(K_2)n_2$, $n_1 \log_a K_1 = n_2 \log_a K_2$, $kK_1/\log_a K_1 = k(K_2)/\log_a K_2$. Из полученного выражения видно, что коэффициент пропорциональности тем больше, чем больше основание используемого кода.

Свяжем количество информации с вероятностью появления отдельных сообщений. Если сообщения равновероятны и на выходе некоторого источника появляется $M$ различных сообщений, то вероятность возникновения каждого сообщения $p(x_{0j}) = 1/M$, $I = -\log_2 p(x_{0j})$. Таким образом, получаем статистическую меру информации, связывающую вероятность появления каждого сообщения и количество информации. Поскольку за основание логарифма принята двоичная единица,

то эта мера представляет собой двоичную единицу на сообщение и отражает количество информации, которое в среднем содержится в каждом равновероятном сообщении. Полученное выражение в общем случае определяет информацию, которая содержится в некотором событии $x_{0j}$ из множества $X_0$ и является функцией ансамбля этого множества. Она всегда неотрицательна и увеличивается с уменьшением вероятности $p(x_{0j})$. Физически данная информация может быть рассмотрена либо как некоторая априорная неопределенность события $x_{0j}$ из множества $X_0$, либо как информация, требуемая для разрешения этой неопределенности. Следует отметить, что данная формула является простейшей: в ней не учитываются некоторые закономерности, связанные с информацией, которая может иметься у наблюдателя до появления данного сообщения, а поэтому весьма существенное место занимает понятие взаимной информации.

Предположим, что на выходе некоторого источника появляется совокупность сообщений из множества $X_0$, которую мы каким-то образом определяем с учетом воздействующих помех посредством ансамбля $Y_0$. Появление некоторого события из ансамбля $Y_0$ изменяет вероятность $p(x_{0j})$ от некоторой априорной вероятности $p(x_{0j})$ до апостериорной вероятности $p(x_{0j}/y_{0j})$. Для оценки количественной меры изменения этой вероятности может быть использован логарифм отношения апостериорной вероятности к априорной, тогда информация о некотором событии из множества $X_0$, содержащаяся в некотором событии из множества $Y_0$:

$$I(x_{0j}, y_{0i}) = \log_2 [p(x_{0j}/y_{0i})/p(x_{0j})].$$

С учетом всех событий, входящих в множества $X_0$, $Y_0$, можно получить окончательно взаимную информацию, как функцию некоторого ансамбля $X_0$, $Y_0$, не зависящую от частных исходов, входящих в эти ансамбли. Суммируя по всем возможным событиям, составляющим ансамбли $X_0$, $Y_0$, получаем

$$I(X_0, Y_0) = \sum_j \sum_i p(x_{0j}, y_{0i}) \times \log_2 [p(x_{0j}/y_{0i})/p(x_{0j})].$$

Нетрудно видеть, что в частном случае, когда появление данного исхода $y_{0j}$ однозначно определяет, что исходом $x_{0j}$ будет некоторый конкретный элемент множества $X$, получаем собственную информацию, содержащуюся в конкретном событии, т. е. в сообщении.

Рассмотренные формулы можно применять для оценки количества информации в реальных условиях передачи. Например, если передается множество двоичных последовательностей длиной $m$ с вероятностью появления каждой последовательности $1/M$, где $M = 2^m$, то собственная информация,

содержащаяся в каждом сообщении, или количество информации в одном усредненном сообщении $I(X_0) = -\log_2 p(x_{0j}) = m$ двоичных единиц, т. е., используя код без избыточности, получаем, что каждый элемент двоичного кода переносит одну двоичную единицу информации. При введении избыточности в код сохраняется число передаваемых сообщений $M$, однако длина кода возрастает до $n$. Количество передаваемой информации составит при равновероятности передаваемых сообщений, как и ранее, $I = \log_2 M$, т. е. $m$ двоичных единиц. Поскольку для передачи $m$ двоичных единиц используется $n$ элементов в коде, где $n > m$, то каждый элемент кода передает $m/n$ двоичных единиц информации, т. е. в одном элементе избыточного кода передается менее одной двоичной единицы информации за счет избыточности, которая тратится либо на обнаружение, либо на обнаружение и исправление ошибок.

Таким образом, аддитивная мера информации позволяет оценить количество информации, передаваемой в одном элементе кода с учетом статистических свойств источника информации, и дает возможность в дальнейшем перейти к оценке скорости передачи информации и сравнению ее с пропускной способностью канала связи, что в целом позволяет дать общую характеристику эффективности использования канала связи, т. е. эффективности согласования источника информации с каналом связи.

## § 2 Количество информации в неравновероятных сообщениях

Количество информации в неравновероятных дискретных сообщениях может быть получено с учетом аддитивности количественной меры информации. Будем считать, что имеется источник сообщений, в котором вероятности появления отдельных сообщений произвольны и в общем случае не подчиняются равномерному закону распределения, т. е. появление сообщений неравновероятно. При выборе каждого конкретного сообщения с некоторой вероятностью $p(x_{0j})$ получаем количество информации, которое должно являться функцией данной вероятности, т. е. $I = \varphi[p(x_{0j})]$.

Пусть за сообщением $x_{0j}$ из источника сообщений появляется сообщение $x_{0k}$, тогда вероятность того, что последовательно возникают сообщения $x_{0j}$, $x_{0k}$, обозначим через $p(x_{0j}, x_{0k})$ и будем считать, что в паре сообщений заключено количество информации, определяемое как $\varphi[p(x_{0j}, x_{0k})]$. Очевидно, что естественным является требование аддитивности получаемой меры информации. Тогда можно считать, что количество информации, заключенное в последовательности двух

сообщений, равно сумме количеств информации, которые содержатся в каждом из выбранных сообщений. Количество информации, содержащееся в сообщении $x_{0j}$, определяется как $\varphi[p(x_{0j})]$. Для оценки количества информации, содержащейся в сообщении $x_{0k}$, необходимо учесть условную вероятность появления сообщения $x_{0k}$ после сообщения $x_{0j}$. Обозначим эту вероятность через $p(x_{0k}/x_{0j})$. Тогда, исходя из условия аддитивности меры информации, можно записать следующее [5]:

$$\varphi[p(x_{0j})] + \varphi[p(x_{0k}/x_{0j})] = \varphi[p(x_{0j}, x_{0k})].$$

Так как $p(x_{0j}, x_{0k}) = p(x_{0j})\, p(x_{0k}/x_{0j})$, то обозначая $p(x_{0j}) = p_j$, $p(x_{0k}/x_{0j}) = p_k$, получаем

$$\varphi(p_j) + \varphi(p_k) = \varphi(p_j, p_k).$$

Для определения функции меры информации $\varphi$ продифференцируем полученное уравнение по величине $p_j$:

$$\varphi'(p_j) = p_k \varphi'(p_j, p_k), \text{ или } p_j\varphi'(p_j) = p_j p_k \varphi'(p_j, p_k).$$

Учитывая нормировку вероятностей, можно записать, что полученное уравнение справедливо при $0 < p_j < 1$, $0 < p_j p_k < p_k$. Последнее условие выполняется, если обе части уравнения представляют собой постоянную величину $k$, т. е. $p_j\varphi'(p_j) = k$. Отсюда $\varphi'(p_j) = k/p_j$. Интегрируя данное уравнение, окончательно получаем

$$\varphi(p_j) = k \ln p_j + C, \text{ или } \varphi[p(x_{0j})] = k \ln [p(x_{0j})] + C.$$

Постоянную $C$ можно определить из условия, что при наличии лишь одного сообщения $x_{0j}$, а этому соответствует $p(x_{0j}) = 1$, количество информации равно нулю, т. е. приход этого сообщения заранее предопределен и никакой неопределенности в получении информации из источника сообщений нет. Поэтому, подставляя $\varphi(1) = 0$ при $p(x_{0j}) = 1$, получаем, что $C = 0$. Тогда

$$\varphi[p(x_{0j})] = k \ln p(x_{0j}).$$

Для выбора коэффициента $k$ необходимо выбрать систему единиц. Если считать, что мера информации должна быть положительной, то наиболее разумным является принятие некоторого отрицательного коэффициента. В простейшем случае при $k = -1$ имеем $\varphi[p(x_{0j})] = -\ln p(x_{0j})$ и за единицу информации принимаем натуральную единицу. Тогда количество информации в одну натуральную единицу равно той информации, которая передается в сообщении, имеющем вероятность появления $1/e$. Следует отметить, что натуральная единица не нашла применения в современной теории связи, чаще

всего используется двоичная единица информации и коэффициент $k$ выбирается в виде $k = -1/\ln 2$, тогда

$$\varphi [p(x_{0j})] = -\ln p(x_{0j})/\ln 2.$$

Окончательно количество информации для неравновероятных сообщений

$$I = \varphi [p(x_{0j})] = -\log_2 p(x_{0j}) = \log_2 [1/p(x_{0j})].$$

Полученная мера информации справедлива для произвольного выбора сообщений из источника, является аддитивной мерой и позволяет измерить количество информации в двоичных единицах.

Пользуясь аддитивностью меры информации, определим количество информации с учетом вероятности появления отдельных элементов кодовой комбинации, отображающей конкретное сообщение. Найдем число сведений, которое содержит некоторое усредненное сообщение, выбранное случайным образом. Этому сообщению соответствует последовательность символов кода $x_{(1)}$, $x_{(2)}$, ..., $x_{(i)}$, ..., $x_{(n)}$. Рассмотрим процесс образования символов, считая, что он описывается цепью Маркова. Символы кода по своему характеру могут иметь $K$ значений. Пусть символ типа $j$ встречается $n_j$ раз, тогда вероятность возникновения сообщения $x_{0j}$ составит

$$p(x_{0j}) = \Pi_{j=1}^{K} [p(x_j)]^{n_j}.$$

Если длина кода велика, то $n_j$ связано с длиной кода соотношением $n_j = np(x_j)$. Так как $I = -\log_2 p(x_{0j})$, то количество информации для неравновероятных сообщений

$$I = -n \sum_{j=1}^{K} p(x_j) \log_2 p(x_j).$$

Из формулы видно, что количество информации прямо пропорционально длине сообщения и зависит от статистики символов кода. Максимально возможное количество информации, которое может быть передано при этом, является максимумом получаемого выражения $I$, что имеет место при $p(x_j) = 1/K$, т. е. когда в одном элементе двоичного кода передается одна двоичная единица информации. Максимум может быть достигнут для случая оптимального построения кода в каналах передачи информации без шума. В каналах с шумом за счет наличия избыточности в одном двоичном элементе кода передается менее одной двоичной единицы информации и количество информации $I < n$.

## § 3 Понятие и свойства энтропии

*Энтропия — мера неопределенности состояния системы.* В информационных системах неопределенность снимается за счет принятой информации, поэтому численно энтропия равна количеству информации, т. е. *энтропия является количественной мерой информации.* При выводе формулы энтропии примем следующие условия [28].

1. Функция энтропии должна быть непрерывной относительно вероятности возникновения сообщений, т. е. можно иметь разное количество информации для одного и того же числа событий, если изменяется закон распределения вероятности их появления. Для дискретных событий наибольшую энтропию, или наибольшее количество информации, дает равномерный закон распределения вероятностей.

2. При одинаковых вероятностях возникновения дискретных событий функция должна быть монотонно возрастающей с увеличением числа событий, т. е. если обеспечен равномерный закон, то для повышения энтропии необходимо увеличивать возможность выбора, например, увеличить основание кода.

3. Функция не должна зависеть от пути выбора событий.

Принимая во внимание перечисленные условия, можно записать

$$H = -k \sum_{j=1}^{K} p(x_j) \log_2 p(x_j),$$

где $x_j$ — символ передаваемого кода, $k$ — коэффициент пропорциональности.

Пусть имеют место два события и $k = 1$, тогда

$$H = -\sum_{j=1}^{2} p(x_j) \log_2 p(x_j).$$

Если меняются вероятности возникновения событий, то оба события составляют полную систему, т. е. $p(x_1) + p(x_2) = 1$. На рис. 13 представлены зависимости $H = f[p(x_1)]$, $H = f[p(x_2)]$.

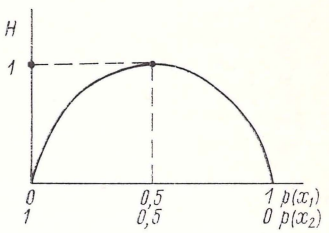

Рис. 13.

Раскроем свойства энтропии. Общая формула энтропии в случае, если имеет место $K$ различных символов $x_j$, при $k = 1$ может быть записана в виде

$$H = -\sum_{j=1}^{K} p(x_j) \log_2 p(x_j).$$

Покажем, что формула удовлетворяет принятым выше условиям.

1. $H \geqslant 0$. Если имеет место единственный символ с вероятностью появления $p(x_j) = 1$, то $H = 0$, т. е. никакой неопределенности нет, выборка из

источника является детерминированной, и никакого количества информации не получаем при появлении единственно возможного символа.

Энтропия положительна, так как вероятности заключены между нулем и единицей. Логарифмы таких чисел отрицательны, следовательно, количество информации, которое получаем из определения энтропии, всегда положительно. При фиксированном $K$ и равновероятности появления всех возможных символов $x_j$ из множества $X$ имеем максимум энтропии, что можно доказать следующим образом. Обозначим $p(x_j) = p_j$, необходимо найти максимум функции

$$H(p_1, p_2, \ldots, p_j, \ldots, p_K) = -\sum_{j=1}^{K} p_j \log_2 p_j.$$

Условием нормировки вероятностей является $\sum_{j=1}^{K} p_j = 1$.

Воспользуемся элементами вариационного исчисления. В общем случае, если имеется некоторая функция $f(x_1, x_2, \ldots, x_k)$, аргументы которой связаны дополнительными условиями

$$\varphi_1(x_1, x_2, \ldots, x_K) = 0, \ \varphi_2(x_1, x_2, \ldots, x_K) = 0, \ \ldots,$$

$$\varphi_n(x_1, x_2, \ldots, x_K) = 0, \ n < K,$$

для нахождения экстремума введем неопределенные множители Лагранжа $\lambda_1, \lambda_2, \ldots, \lambda_n$ и составим функцию

$$\Phi(x_1, x_2, \ldots, x_K, \lambda_1, \lambda_2, \ldots, \lambda_n) = f(x_1, x_2, \ldots, x_K) + \\ + \lambda_1 \varphi_1(x_1, x_2, \ldots, x_K) + \ldots + \lambda_n \varphi_n(x_1, x_2, \ldots, x_K).$$

Продифференцируем $\Phi$ по всем аргументам последовательно и приравняем частные производные нулю. Тогда получим необходимые условия для экстремума:

$$\frac{\partial \Phi(x_1, x_2, \ldots, x_K)}{\partial x_1} = 0, \ \ldots, \ \frac{\partial \Phi(x_1, x_2, \ldots, x_K)}{\partial x_K} = 0,$$

$$\frac{\partial \Phi(x_1, x_2, \ldots, x_K)}{\partial \lambda_1} = 0, \ \ldots, \ \frac{\partial \Phi(x_1, x_2, \ldots, x_K)}{\partial \lambda_n} = 0.$$

Решая систему $n + K$ уравнений, найдем значения аргументов, при которых наступает экстремум. Для определения характера экстремума необходимо найти знак второй производной.

Следуя данной методике, в нашем случае обозначим

$$\Phi(p_1, p_2, \ldots, p_K) = H(p_1, p_2, \ldots, p_K) + \lambda\left(\sum_{j=1}^{K} p_j - 1\right),$$

где $\lambda$ — неопределенный коэффициент. Производные от функции $\Phi$ по вероятностям $p_j$ приравняем нулю:

$$\partial \left\{[H(p_1, p_2, \ldots, p_K) + \lambda\left[\sum_{j=1}^{K} p_j - 1\right]\right\}/\partial p_j = 0.$$

Подставляя значение $H(p_1, p_2,\ldots,p_k)$, получаем

$$\partial\left[-\sum_{j=1}^{K} p_j \log_2 p_j + \lambda\left(\sum_{j=1}^{K} p_j - 1\right)\right]/\partial p_j = 0.$$

Отсюда $-(\ln p_j + 1)/\ln 2 + \lambda = 0$, $p_j = \exp(\lambda \ln 2 - 1)$. Учтем условие $\sum_{j=1}^{K} p_j = 1$:

$$\sum_{j=1}^{K} \exp(\lambda \ln 2 - 1) = K \exp(\lambda \ln 2 - 1) = 1,$$

$$\exp(\lambda \ln 2 - 1) = 1/K = p_j.$$

Итак, при $p_j = 1/K$ получаем экстремум функции энтропии. Для определения характера экстремума найдем вторую производную:

$$\partial^2\left[H(p_1, p_2, \ldots, p_K) + \lambda\left(\sum_{j=1}^{K} p_j - 1\right)\right]/\partial p_j^2 = -1/p_j \ln 2 < 0.$$

Так как она отрицательна, то значение $p_j = p(x_j) = 1/K$ соответствует максимуму функции, т. е. энтропия имеет максимум при условии равновероятности появления символов, или равновероятности дискретных событий. Максимум может быть найден при $p_j = 1/K$:

$$H = -(K/K)\log_2(1/K) = \log_2 K.$$

Таким образом, в случае равновероятности входных событий энтропия соответствует количеству информации для равновероятных исходов. Действительно, если рассматривать $M$ событий с вероятностью появления каждого сообщения $p(x_{0j}) = 1/M$, то энтропия $H = \log_2 M$. Полученное выражение соответствует формуле Хартли для равновероятных сообщений и определяет количество информации, содержащееся в любом сообщении при равновероятности их появления, т. е. формула Хартли соответствует максимальному значению энтропии. Физически это определяет случай, когда неопределенность настолько велика, что прогнозировать оказывается трудно.

2. Второе свойство энтропии соответствует случаю равномерного закона появления отдельных событий. Будем считать, что определяем энтропию как количество информации, содержащееся в сообщении, а множество сообщений $M$ увеличивается, тогда $H(X_0) = \log_2 M$.

Нетрудно видеть, что при возрастании $M$ функция $H(X_0)$ является монотонно возрастающей, а поэтому при увеличении возможности выбора неопределенность возрастает, и большую долю информации вносит каждый последующий выбор.

3. Функция энтропии при наличии различных сложных опытов, объединяющих ряд событий, не должна зависеть от пути выбора события. Рассмотрим некоторый сложный

опыт, объединяющий опыт $X$, включающий в себя $n$ событий вида $x_1, x_2, \ldots, x_n$ с вероятностями наступления $p(x_1), p(x_2), \ldots, p(x_n)$, и опыт $Y$, включающий в себя ряд событий $y_1, y_2, \ldots, y_m$ с вероятностями наступления этих событий $p(y_1), p(y_2), \ldots, p(y_m)$, Можно рассмотреть все возможные совместные исходы $x_j$, $y_i$ и установить вероятности появления пар исходов $p(x_j, y_i)$. При этом могут возникнуть вероятности появления пар исходов $p(x_1, y_1) \ldots, p(x_n, y_m)$. Энтропия произвольного сложного опыта может быть определена как

$$H(X, Y) = -\sum_{j=1}^{n} \sum_{i=1}^{m} p(x_j, y_i) \log_2 p(x_j, y_i).$$

Расчет необходимо произвести при условии

$$\sum_{j=1}^{n} \sum_{i=1}^{m} p(x_j, y_i) = 1.$$

Независимо от последовательности проведения этих опытов получаем одну и ту же величину, что свидетельствует о неопределенности в совместном наступлении событий $x_j$, $y_i$, имеющей место до опыта.

Описанный выше случай можно рассматривать как пример объединения совокупностей двух ансамблей случайных событий. Объединения могут быть различными, в соответствии с чем различаются и некоторые виды энтропии. Для ансамблей событий $X$, $Y$ можно выделить безусловную энтропию ансамбля $H(X)$, показывающую среднее количество информации на один элемент ансамбля $X$. Соответственно для ансамбля $Y$ получаем безусловную энтропию $H(Y)$. Можно выделить взаимную энтропию $H(X, Y)$, рассматривая сложный опыт, включающий в себя множество событий $X$, $Y$. Эта энтропия характеризует среднюю информацию, которая приходится на пару символов, выбранных из множества $X$, $Y$.

Существенное значение в теории передачи информации и различных физических явлениях имеет условная энтропия. Если символы в коде являются взаимно зависимыми, т. е. последующий символ связан статистической зависимостью с предыдущим, то количество информации, содержащееся в символе, уменьшается, и это оценивается условной энтропией.

Пусть имеется символ $x_j$, за которым следует символ $y_i$, тогда $p(y_i/x_j)$ — вероятность появления $y_i$ при наличии $x_j$. Найдем количество информации, которую несет символ $y_i$:

$$H(y_i/x_j) = -p(x_j)\, p(y_i/x_i) \log_2 p(y_i/x_i).$$

Количество информации в символе $y_i$ при условии, что перед ним имел место любой символ из множества $X$:

$$H(y_i/X) = -\sum_{j=1}^{K} p(x_j)\, p(y_i/x_j) \log_2 p(y_i/x_j).$$

Предположим, что символы $y_i$ составляют множество $Y$, тогда условная энтропия между множествами $Y$ и $X$

$$H(Y/X) = -\sum_{j=1}^{K}\sum_{i=1}^{K} p(x_j)\, p(y_i/x_j) \log_2 p(y_i/x_j).$$

Условная энтропия используется для оценки количества информации, которую вносит помеха в сигнал, передаваемый по каналу связи.

Распространим понятие энтропии на сообщения, имеющие непрерывный характер. Положим, что непрерывное сообщение обладает известной плотностью распределения вероятности $W(x)$, которая характеризует вероятность попадания непрерывной величины $x$ в интервал $dx$. Непрерывное сообщение характеризуется бесконечно большим набором элементов $x_j$. Выберем эти элементы с постоянным интервалом $\Delta x$ и положим, что число их конечно и элементы имеют вид $x_1, x_2, \ldots, x_j, \ldots, x_K$. Тогда можно использовать введенное выше выражение энтропии

$$H = -\sum_{j=1}^{K} p(x_j) \log_2 p(x_j).$$

Для элемента $x_j$ вероятность появления $p(x_j)$ можно найти из распределения $W(x)$ при $x = x_j$ с учетом зоны дискретизации элемента $\Delta x$ в виде [7] $p(x_j) = W(x_j)\,\Delta x$. Отсюда

$$H = -\sum_{j=1}^{K} W(x_j)\,\Delta x \log_2 W(x_j) - \sum_{j=1}^{K} W(x_j)\,\Delta x \log_2 \Delta x.$$

Так как сообщение имеет непрерывный характер, то $\Delta x \to 0$, $K \to \infty$. Тогда

$$H = -\int_{-\infty}^{\infty} W(x) \log_2 W(x)\, dx - \int_{-\infty}^{\infty} W(x) \log_2 \Delta x dx =$$

$$= -\int W(x) \log_2 [W(x)\,\Delta x]\, dx.$$

При $\Delta x \to 0$, учитывая, что $\int_{-\infty}^{\infty} W(x)\, dx = 1$, получаем

$$H = -\int_{-\infty}^{\infty} W(x) \log_2 W(x)\, dx - \log_2 \Delta x.$$

Так как количество информации определяется в основном первым слагаемым, то зачастую рассчитывают приведенную относительную энтропию

$$H_{\text{п}} = -\int_{-\infty}^{\infty} W(x) \log_2 W(x)\, dx.$$

При этом следует иметь в виду, что

$$H = -\int_{-\infty}^{\infty} W(x) \log_2 [W(x)\,\Delta x]\, dx.$$

Очевидно, что численное значение энтропии непрерывной величины определяется видом распределения $W(x)$. Так же

как и для дискретных сообщений, целесообразно в данном случае установить вид оптимального распределения $W_{\text{опт}}(x)$, при котором $H$ максимально. Анализ выполним для случая, когда дисперсия элементов сообщений есть постоянная величина. Тогда вид оптимального распределения найдем при следующих ограничениях:

$$\int_{-\infty}^{\infty} W(x)\,dx = 1,\ \int_{-\infty}^{\infty} x^2 W(x)\,dx = \sigma^2,$$

где $\sigma$ — среднеквадратичное отклонение от математического ожидания $\bar{x} = 0$.

Аналогично предыдущему на основе вариационного исчисления введем функцию

$$\Phi(x) = f(x) + \lambda_1\varphi_1(x) + \lambda_2\varphi_2(x),$$

где $\lambda_1$, $\lambda_2$ — неопределенные множители Лагранжа.

Найдем производную от функции $\Phi(x)$ по $W(x)$ и приравняем ее нулю:

$$\partial\Phi(x)/\partial W(x) = -\log_2[W(x)\,\Delta x] - \log_2 e + \lambda_1 + \lambda_2 x^2 = 0.$$

Отсюда $\log_2[eW(x)\,\Delta x] = \lambda_1 + \lambda_2 x^2$.

Оптимальное распределение

$$W_{\text{опт}}(x) = \exp(\lambda_1 \ln 2 - 1)\exp(\lambda_2 \ln 2x^2)/\Delta x.$$

Учитывая условия

$$\int_{-\infty}^{\infty}\left[\frac{1}{\Delta x}\exp(\lambda_1 \ln 2 + \lambda_2 \ln 2x^2 - 1)\right]dx = 1,$$

$$\int_{-\infty}^{\infty}\left[\frac{x^2}{\Delta x}\exp(\lambda_1 \ln 2 + \lambda_2 x^2 - 1)\right]dx = \sigma^2,$$

находим $W_{\text{опт}}(x) = \left[1/\sqrt{2\pi\sigma^2}\right]\exp(-x^2/2\sigma^2)$.

Таким образом, при $\bar{x} = 0$ и ограниченной дисперсии оптимальным распределением $W(x)$, обеспечивающим максимум энтропии, является нормальное распределение. Этот результат используется практически при передаче непрерывной функции в виде отклонений от математического ожидания, что повышает количество передаваемой информации через канал связи. Найдем численное значение максимума энтропии:

$$H_{\max} = -\int_{-\infty}^{\infty} W_{\text{опт}}(x)\log_2 W_{\text{опт}}(x)\,dx - \log_2\Delta x =$$

$$= -\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\log_2\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\right]dx -$$

$$-\log_2\Delta x = \log_2\frac{\sigma\sqrt{2\pi e}}{\Delta x}.$$

Соответственно максимум приведенной энтропии

$$H_{\text{п max}} = \log_2 [\sigma \sqrt{2\pi e}\,].$$

К непрерывным сообщениям приложимо и понятие условной энтропии. Если имеются сообщения, характеризующиеся непрерывным распределением элементов $W(x)$, и зависимые от них сообщения с распределением элементов $W(y/x)$, то, применяя к каждому бесконечному множеству элементов $x$ и $y$ фиксацию элементов с интервалами $\Delta x$, $\Delta y$, получаем конечные множества $x_1, x_2, \ldots, x_K$, $y_1, y_2, \ldots, y_K$.

При $\Delta x \to 0$, $\Delta y \to 0$, $K \to \infty$ находим условную энтропию в виде

$$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x)\, W(y/x) \log_2 W(y/x)\, dxdy.$$

Условная энтропия позволяет учесть потери непрерывной информации при ее передаче по каналу связи с помехами.

## § 4. Энтропия источника и энтропия сообщения

Пусть имеется некоторый отправитель (О), который выдает дискретные сообщения $X_0$. С помощью кодирующего устройства (КУ) каждое сообщение превращается в код. Множество символов кода обозначим через $X$. Если исследуется канал связи (КС) (рис. 14), то можно не обращаться к отправителю, а рассматривать лишь источник символов (кодирующее устройство). Тогда возникает необходимость связать свойства источника и отправителя. Эта связь возможна через энтропию.

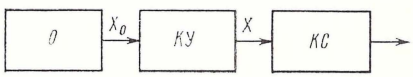


Рис. 14.

Под энтропией сообщения будем понимать количество информации, содержащееся в любом усредненном сообщении. Тогда

$$H(X_0) = -\sum_{j=1}^{M} p(x_{0j}) \log_2 p(x_{0j}) \left[\frac{\text{двоичных единиц}}{\text{сообщение}}\right]$$

— усредненная энтропия сообщения. Соответственно энтропия источника, или количество информации, содержащееся в одном символе сообщения:

$$H(X) = -\sum_{j=1}^{K} p(x_j) \log_2 p(x_j) \left[\frac{\text{двоичных единиц}}{\text{символ}}\right].$$

Пример. Пусть передается четыре равновероятных сообщения двоичным неизбыточным кодом. Сообщения отображаются кодом 00, 01, 10, 11.

Найдем энтропию сообщения $H(X_0) = -\sum_4 (1/4) \log_2 (1/4) = 2$ и энтропию источника $H(X) = -\sum_2 (1/2) \log_2 (1/2) = 1$.

Из примера видно, что каждый символ несет одну двоичную единицу информации.

Разделим $H(X_0)$ на $H(X)$ и получим число элементов в коде, т. е. $H(X_0)/H(X) = n$. Если данное условие соблюдается, то код называется оптимальным, в противном случае в коде возникает избыточность и он становится неоптимальным для канала без шума [21]. Для получения оптимального кода необходимо, чтобы символы в нем встречались с равной вероятностью.

Полученные соотношения для энтропии источника и энтропии сообщения ограничены случаем взаимной независимости появления либо отдельных символов источника, либо отдельных сообщений отправителя. В ряде практических случаев встречается более общий класс источников, где вероятность выбора одного символа зависит от того, какие символы выбирались до этого. Зачастую можно ввести ограничение, указывающее длительность корреляционных связей во времени. В тех случаях, когда можно указать достаточно большой отрезок времени, на котором связи не распространяются, целесообразно воспользоваться математическим представлением источника в виде цепи Маркова. Цепь Маркова обычно характеризует последовательность событий, при которых вероятность появления данного события зависит от исходов, имеющих место в $n$ предшествующих событиях. Тогда получаем цепь Маркова порядка $n$ и можем не учитывать вероятности, которые связывают $n+1$, $n+2$, ... событий с данным событием. Однако, учитывая интервал $n$ предшествующих событий, можно указать, что появление данного события в значительной степени зависит от того, какие конкретно $n$ событий или $n$ символов возникли до этого. Поэтому можно ввести понятие состояния источника с указанием вероятности возникновения данного состояния.

Чтобы охарактеризовать источник с взаимно зависимым появлением отдельных символов, необходимо знать вероятности, определяющие все множество состояний и для каждого $q$-го состояния найти вероятность $P_q$ появления этого состояния. Тогда могут быть найдены вероятности $p_q(x_k)$ возникновения любого $k$-го символа в $q$-м состоянии источника. Среднее количество информации в символе для источника, находящегося в $q$-м состоянии, можно назвать энтропией этого состояния и определить для данного состояния $q$:

$$H_q = -\sum_{k=1}^{K} p_q(x_k) \log_2 p_q(x_k).$$

Если возможно появление $K$ различных символов и $l$ состояний источника, то энтропия источника

$$H(X) = -\sum_{q=1}^{l} \sum_{k=1}^{K} P_q p_q(x_k) \log_2 p_q(x_k).$$

Следует отметить, что наличие вероятностных связей в источнике вызывает уменьшение энтропии источника.

Для характеристики источника обычно вводят понятие избыточности. При этом сравнивают действительную энтропию с максимально возможной. Очевидно, что максимально возможной энтропии источника соответствует максимум информации, которая может передаваться в одном символе. Максимум отвечает равномерному распределению вероятностей появления дискретных символов, т. е. существует при $p(x_j) = 1/K$. В этом случае $H_{\max}(X) = \log_2 K$. Избыточность некоторого источника можно определить как $И = 1 - H(X)/H_{\max}(X)$. Отсюда

$$И = 1 - H(X)/\log_2 K.$$

Наличие избыточности в источнике связано, с одной стороны, с вероятностными зависимостями между соседними символами и, с другой — с неравномерным законом распределения вероятностей появления отдельных символов из источника. Для увеличения энтропии источника необходимо стремиться строить источник с взаимно независимыми символами при равновероятности появления каждого символа из источника.

Совокупность символов, выдаваемых из источника, составляет некоторое сообщение. При этом в случае использования неизбыточных кодов число возможных сообщений $M = К^n$. Если каждое сообщение включает в себя $n$ элементов, то из-за аддитивности меры информации получаем, что количество информации, передаваемой одним сообщением, складывается из количеств информации, передаваемой всеми элементами кода. Тогда в простейшем случае неизбыточного кода получаем $H(X_0) = nH(X)$. Избыточность, возникающая в каждом символе кода, пропорционально распределяется и на сообщения. Если максимум информации, которую может нести сообщение, составляет $H_{\max}(X_0) = \log_2 M$, то очевидно, что он возникает при наличии максимума информации в каждом символе кода, а поэтому $H_{\max}(X_0) = nH_{\max}(X) = n\log_2 K$. Тогда избыточность сообщения

$$И_с = 1 - H(X_0)/H_{\max}(X_0) =$$
$$= 1 - nH(X)/nH_{\max}(X) = И.$$

Таким образом, избыточность при изменении алфавита не меняется, однако с увеличением длины сообщения возникает эффект декорреляции отдельных символов и происходит устранение взаимных связей между ними. Для источника без избыточности энтропия источника является максимальной и равна $\log_2 K$. Если данный источник формирует равновероятные взаимно независимые последовательности длиной $n$ так,

что $M_{\min} = K^n$, то такие последовательности являются типичными, и для любой последовательности вероятность появления $p(x_{0j}) = 1/K^n$. Отсюда $H_{\max}(X) = \log_2 p(x_{0j})/n$. При этом общее число типичных последовательностей длиной $n$ с увеличением длительности последовательностей асимптотически приближается к величине $M_{\min} = 2^{-nH_{\max}(X)}$. Физически типичная последовательность характеризует оптимальное использование каждого символа в коде, и среди всех возможных последовательностей соотношение типичных последовательностей может быть определено как

$$M_{\min}/M = 2^{-n[H_{\max}(X)-H(X)]}.$$

Очевидно, что при достаточно большом $n$ $H(X)$ приближается к $H_{\max}(X)$, и все возможные последовательности будут типичными.

## § 5. Связь количества информации с эффективностью управления

В основе функционирования автоматизированной системы управления лежит заранее определенная цепь управления. Отклонение выходного состояния системы, характеризуемое некоторым выходным вектором относительно вектора цели, приводит к снижению эффективности функционирования, и степень этого отклонения определяет рассогласование в процессе функционирования системы. Уменьшение рассогласования достигается путем выработки управляющей информации, которая возникает на основе обработки осведомляющей информации, характеризующей процесс производства. Наличие различных этапов преобразования информации приводит к ее искажению вследствие снижения качества функционирования комплекса технических средств АСУ, даже если выработана оптимальная по управлению информация. Искажения возникают в результате того, что между выходным вектором системы и вектором цели на разных этапах преобразования информации образуется рассогласование.

Понятие энтропии позволяет охарактеризовать функционирование системы с позиции целевой неупорядоченности в системе. Так, целевую энтропию, т. е. неупорядоченность в управляющей информации, можно определить в следующем виде. Если в системе имеется $r$ этапов преобразования информации и возникает рассогласование выходного вектора системы относительно вектора цели на $i$-м этапе в виде $q_i$, то

$$H_{\text{у}} = \sum_{i=1}^{r} p_i \log_2 [1/(q_i + 1)],$$

где $p_i$ — вероятность появления рассогласования величиной $q_i$. Характерно, что $\sum_{i=1}^{r} p_i = 1$. Степень рассогласования в виде

целевой энтропии характеризуется в двоичных единицах информации.

В зависимости от метода измерения рассогласования энтропия может быть непрерывной или квантованной, ее можно оценить либо по сравнению с постоянным рассогласованием, либо по отношению к некоторой области. Таким образом, функция целевой энтропии позволяет определить меру неупорядоченности функционирования системы в отношении достижения поставленной цели управления. Для определения ценности управляющей информации нужно рассмотреть разность целевых энтропий до и после замыкания канала управления на объект управления. Тогда количество информации $I_{\text{у}} = H_{\text{у.п}} - H_{\text{у.д}}$.

Представляет интерес связать количество информации, необходимой для управления, с эффективностью управления. Будем считать, что при идеальном управлении достигается максимум эффективности $\text{э}_{\max}$. Тогда если ввести некоторую величину $N$, характеризующую разнообразие системы при данном способе управления, то эффективность реального способа управления конкретной АСУ может быть оценена как $\text{э} = \text{э}_{\max}(1 - N/N_0)$, где $N/N_0 < 1$. Очевидно, что при управлении необходимо уменьшить число возможных состояний, т. е. уменьшить ее неупорядоченность. Можно ввести функцию энтропии $H = k \log_2 N$, где $k$ — некоторый постоянный коэффициент. Чтобы поднять эффективность управления, необходимо уменьшить энтропию, поэтому можно записать, что количество управляющей информации $I_{\text{у}} = H_0 - H = k \log_2 (N_0/N)$, а $N = N_0 \exp(-I/k)$. Энтропия $H_0$ соответствует некоторой исходной неупорядоченности $N_0$ в системе. Учитывая формулу для эффективности, получаем $\text{э} = \text{э}_{\max}\{1 - \exp(-I/k)\}$. Отсюда видно, что вследствие экспоненциальной зависимости между эффективностью и количеством необходимой для управления информации $I_{\text{у}}$ весьма существенным является наличие некоторой начальной доли управляющей информации. Непрерывное увеличение управляющей информации оказывается нецелесообразным, т. е. в системе необходимо сохранить некоторую долю неопределенности.

Подробная детализация процесса управления не дает значительного дополнительного экономического эффекта. Если перейти к денежному выражению и считать, что стоимость системы $C = bI_{\text{у}}$, где $b$ — коэффициент пропорциональности, то $\text{э} = \text{э}_{\max}[1 - \exp(-C/bk)]$. Вводя срок окупаемости $t_{\text{ок}} = \Delta C/\Delta\text{э}$, можно определить, что $t_{\text{ок}} = bk/N$. Таким образом, для уменьшения срока окупаемости не следует усложнять управление, так как дополнительное усложнение управления приводит к значительному возрастанию его. Отсюда можно сделать вывод, что не следует стремиться усложнять алгоритмы управления в системах. Зачастую более простые при-

ближенные алгоритмы могут дать уже достаточный экономический эффект.

Наибольший эффект от управления может быть получен в иерархических системах, где на каждом уровне управления можно оставить некоторую долю неупорядоченности и вследствие этого не использовать сложные алгоритмы. Разумеется, возникающие при этом отклонения эффективности от максимально возможной должны быть скомпенсированы соответствующими управляющими воздействиями, идущими с верхнего уровня. Частота этих управляющих воздействий должна быть достаточно низкой, и на нижних уровнях иерархии системы в значительной степени должна быть обеспечена автоматизация процесса управления.

Таким образом, количество управляющей информации в системе является определяющим фактором, от которого зависит эффективность функционирования АСУ. Снижение количества управляющей информации приводит к снижению эффективности, а поэтому целью проектирования КТС АСУ и создания информационного и математического обеспечения является такое построение алгоритмов и структуры комплекса, при котором потери информации на каждом этапе ее преобразования будут минимальными. Следует отметить, что в настоящее время весьма серьезные задачи возникают на таких фазах преобразования информации, как сбор и передача информации, где до сих пор отсутствуют совершенные технические средства, позволяющие преобразовывать информацию с требуемыми вероятностно-временными характеристиками. В соответствии с этим информационный подход к проектированию технического и информационного обеспечения АСУ является исключительно важным.

## *Глава III.*

# СКОРОСТЬ ПЕРЕДАЧИ ИНФОРМАЦИИ И ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛОВ СВЯЗИ АСУ

## § 1. Понятие канала связи. Характеристики каналов связи АСУ

В автоматизированных системах управления каналы связи объединяются в сложную информационную сеть. Однако на каждом уровне иерархии информационной сети осуществляется построение канала связи, согласованного с источником информации, а поэтому основные проблемы передачи информации по непрерывным и дискретным каналам связи являются основополагающими для проектирования оптимальных систем передачи информации в АСУ.

В зависимости от типа информации, которая передается в АСУ, канал связи представляет собой совокупность технических средств, объединенных с линией связи, которые позволяют с заданной степенью верности обеспечить передачу требуемого вида информации.

В настоящее время в АСУ получили широкое применение как непрерывная, так и дискретная информация. В соответствии с этим различают два основных вида каналов связи: непрерывные и дискретные. Непрерывные каналы предназначены для передачи сигналов, представляющих собой непрерывные функции времени, дискретные — для передачи дискретных сигналов и зачастую являются математическим отображением реальных каналов.

В конкретной системе передачи информации обычно присутствуют и дискретный, и непрерывный каналы. Дискретный канал включает в себя модулятор, предназначенный для модуляции исходного переносчика информации некоторым управляющим воздействием, в качестве которого выступает элемент кода или некоторая непрерывная информация. Из модулятора информация через согласующее линейное устройство поступает в непрерывный канал, для которого известны физические характеристики, позволяющие обобщенно определить канал связи. Из непрерывного канала информация через линейное согла-

сующее устройство поступает на демодулятор, позволяющий выделить огибающую передаваемого сигнала. В зависимости от назначения огибающая может представлять собой либо исходную непрерывную функцию, подлежащую передаче, либо управляющий сигнал в виде элемента кода, если передача идет в кодированном виде.

Независимо от использования и назначения непрерывного канала основными его характеристиками являются физические характеристики [24], позволяющие обобщенно оценить возможность использования непрерывного канала для передачи информации: время, на которое канал связи предоставлен для передачи сигнала, $T_к$; ширина полосы пропускания канала связи $F_к$; допустимое превышение сигнала над помехой в данном канале $H_к$.

Перечисленные характеристики определяются физической структурой используемого канала связи и зависят от его типа. Прежде всего от типа канала зависит ширина пропускания. В настоящее время выделяют следующие каналы связи: проводные, акустические, оптические. Ширина спектра для этих каналов меняется от 3300 Гц (для проводных линий телефонного типа) до сотен мегагерц (для оптических линий связи). При использовании различных видов модуляции меняется спектр сигнала: обычно он расширяется. Так, например, при амплитудной модуляции ширина спектра возрастает в два раза, при частотной модуляции — в $2\beta$ раз, где $\beta$ — индекс частотной модуляции.

Временной параметр канала $T_к$ зависит от способа использования канала связи для передачи многих сообщений. Зачастую в системах используется временное уплотнение каналов, и в этом случае для передачи каждой конкретной информации отводится свое временное окно с длительностью $T_к$, в котором разрешается передавать данную информацию, т. е. последовательно во времени по одному и тому же физическому каналу связи осуществляется передача информации разного назначения. Допустимое превышение сигнала над помехой в канале определяется в основном физическими характеристиками канала связи: для проводных каналов — уровнем перекрестных помех и пробивным напряжением изоляции, для радиоканалов — возможностями модуляции и выявления сигнала на соответствующих расстояниях.

Физические характеристики сигнала позволяют в обобщенном виде оценить возможность использования непрерывного канала для передачи сигнала. В соответствии с этим рассмотрим вопрос о выборе характеристик сигнала, который может быть передан по каналу с заданными физическими характеристиками.

Любой передаваемый сигнал, отображающий данное сообщение, можно характеризовать следующими физическими пара-

метрами: длительностью сигнала $T$, спектром сигнала $F$, превышением сигнала над помехой $H$ [24]. Длительность сигнала есть время существования сигнала в канале связи. Она зависит от необходимого быстродействия системы и может составлять для одного сообщения 10 мкс — 1 с.

Ширина импульса сигнала определяется исходя из спектральных свойств. Если сигнал описывается функцией $f(t)$, то плотность амплитудного спектра

$$s(\omega) = \int_T f(t) \exp(-i\omega t)\, dt,$$

соответственно $f(t) = (1/2\pi) \int_F s(\omega) \exp(i\omega t)\, d\omega$.

Рассмотрим плотность спектра одиночного импульса с амплитудой $A$, длительностью $\tau$:

$$s(\omega) = \int_{-\tau/2}^{\tau/2} A \exp(-i\omega t)\, dt = (2A/\omega) \sin(\omega\tau/2) =$$
$$= [A\tau \sin(\omega\tau/2)]/(\omega\tau/2).$$

Для $\sin(\omega\tau/2) = 0$ $\omega = 2n\pi/\tau$. Если ограничиться первой гармоникой, то $2\pi F = 2\pi/\tau$, отсюда $F = 1/\tau$ (рис. 15). В теории связи принято, что $F = \mu/\tau$, где $\mu = 1 \div 5$.

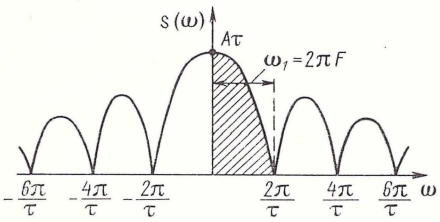


Рис. 15.

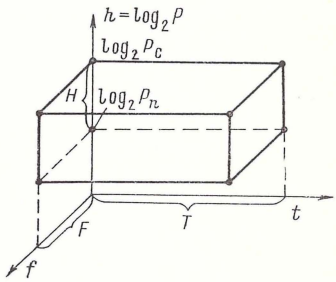


Рис. 16.

Превышение сигнала над помехой определяется возможностями передатчика и дальностью передачи. В промышленных системах превышение сигнала над помехой определяется по формуле

$$H = \log_2 (P_c/P_п),$$

где $P_c$ — мощность сигнала на входе приемника, $P_п$ — мощность помехи на входе приемника.

Введем понятие объема передаваемого сигнала, тогда $V = TFH$. На рис. 16 представлен объем сигнала. Для оценки возможности передачи сигнала по данному каналу введем понятие объема канала $V_к = T_к F_к H_к$. Выполнение неравенства $V \leqslant V_к$ является необходимым условием возможности передачи сигнала по данному каналу, т. е. в этом случае принципиально возможна передача заданного объема сигнала по каналу.

Для непосредственной реализации такой возможности необходимо выполнение достаточных условий: $T \leqslant T_к$, $F \leqslant F_к$, $H \leqslant H_к$. Однако, как правило, все достаточные условия не выполняются, но при выполнении необходимого условия достаточное можно выполнить путем преобразования объема сигнала.

Как было показано выше, произведение длительности сигнала на ширину спектра является постоянной величиной, а поэтому можно за счет взаимного обмена между шириной спектра и длительностью выполнить два условия. Во-первых, если не удовлетворяется условие по превышению сигнала, то можно преобразовать сигнал путем многократной передачи его во времени при пониженном уровне, т. е. произвести взаимный обмен между превышением сигнала и длительностью. Снижение превышения приводит к повышению длительности передачи сигнала. Во-вторых, изменить превышение, т. е. уменьшить его за счет ширины спектра либо за счет длительности, можно путем преобразования сигнала кодированием. Увеличение основания системы счисления кода вызывает расширение спектра сигнала, изменение длительности его и уменьшение необходимого уровня его мощности, т. е. позволяет решить вопрос о снижении превышения сигнала над помехой.

Таким образом, в большинстве случаев при выполнении необходимого условия удается оптимально использовать непрерывный канал и тем самым осуществить передачу заданного объема информации.

Непрерывный канал в общем случае является составной частью дискретного канала. Если на его передающей стороне ввести модулятор, а на приемной стороне — демодулятор, то получим дискретный канал, на входе которого появляется множество символов $X$, а на выходе — множество символов $Y$. Конкретное содержание передаваемых сообщений для канала несущественно, так как канал реагирует в основном на отдельные элементы кода как на управляющие воздействия для модулятора. Множества $X$ и $Y$ являются тождественными, если в канале отсутствуют помехи.

Модель дискретного канала связи обычно изображают в виде графа, в узлах которого находятся формируемые и выявляемые символы, а дуги или ребра отображают вероятности перехода одного символа в другой при передаче его по дискретному каналу связи [22]. На рис. 17 представлен граф модели дискретного канала связи в случае, когда множество символов $X$ на входе имеет размерность $K_x$, а множество символов $Y$ на выходе имеет размерность $K_y$.

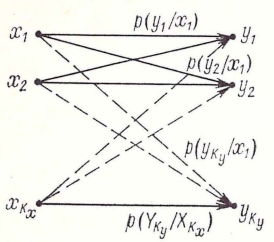


Рис. 17.

Вероятности переходов, связываю-

щие входные и выходные символы в дискретном канале связи, могут быть записаны в виде матрицы

$$P = \begin{Vmatrix} p(y_1/x_1) & p(y_2/x_1) & \dots & p(y_{K_y}/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & \dots & p(y_{K_y}/x_2) \\ \dots & \dots & \dots & \dots \\ p(y_1/x_{K_x}) & p(y_2/x_{K_x}) & \dots & p(y_{K_y}/x_{K_x}) \end{Vmatrix}.$$

Вероятности, расположенные по диагонали, характеризуют собой правильность прохождения символов и по отношению к другим должны иметь наибольшие значения. Сумма всех вероятностей, входящих в любую строку или любой столбец данной матрицы, равна единице.

Рассматривая дискретное отображение реального канала связи, можно выделить следующие основные его характеристики. Каналы можно характеризовать по количеству возможных символов, поступающих на вход и на выход. При этом в зависимости от значения $K_x$ различают двоичные, троичные, четверичные и т. д. каналы связи. Если $K_y > K_x$, то канал называют каналом со стиранием: на приемной стороне появляется некоторая дополнительная область стирания. В этом случае порог срабатывания, разделявший ранее зоны приема единиц и нулей, превращается в зону неопределенности. Попадание символа в зону неопределенности означает неясность в выявлении символа, и это может быть использовано как оценка места ошибки в принятой кодовой комбинации. Введение зоны стирания, разделяющей зоны приема единиц и нулей, позволяет повысить возможность коррекции ошибок и при том же кодовом расстоянии становится возможным кроме обнаружения исправлять ошибки.

Двоичный канал со стиранием используется при исправлении одиночной ошибки.

Вероятности, входящие в матрицу канала $P$, могут быть функциями времени, что характерно практически для всех реальных каналов связи. Наличие зависимости вероятностей от времени приводит к нестационарному каналу связи, характеристики которого являются переменными. Для оценки нестационарных каналов в настоящее время используются модели стационарных каналов, т. е. нестационарный канал рассматривается как некоторая последовательность стационарных, оценивается множество возможных аппроксимирующих стационарных состояний и указываются вероятности попадания канала в каждое аппроксимирующее состояние.

Следует отметить, что в ряде случаев конкретные каналы связи могут быть рассмотрены как каналы стационарные, т. е. каналы, для которых данные вероятности являются постоянными

величинами. Вероятности, входящие в матрицу канала, могут быть зависимы относительно предшествующих значений символов. Тогда возникает модель канала с памятью. Для рассмотрения и оценки качества передачи информации по каналам с памятью используется математический аппарат цепей Маркова, и такие каналы называются марковскими. Для марковского канала характерно существование некоторого конечного множества состояний, для каждого из которых необходимо указать соответствующую матрицу. Тогда число возможных матриц канала с памятью характеризуется числом выбранных состояний канала и могут быть найдены вероятности, усредненные по всем возможным состояниям. Так, если $S$ — число возможных состояний канала, то условная вероятность

$$p(y_i/x_i) = \sum_{k=1}^{S} P_k p_k (y_i/x_j),$$

где $p_k(y_i/x_j)$ — вероятность выявления символа $y_i$ при условии, что был передан символ $x_j$ и канал находился в $k$-м состоянии; $P_k$ — вероятность нахождения канала в $k$-м состоянии.

Дополнительные упрощающие условия могут быть получены при рассмотрении вероятностей, входящих в строки и столбцы матрицы $P$. Если вероятности, входящие в строку матрицы, являются перестановками одних и тех же чисел, то канал называется симметричным по входу. Если это же относится к вероятностям, входящим в столбцы, то канал является симметричным по выходу. При соблюдении обоих предыдущих условий канал является симметричным по входу и по выходу. Модель симметричного дискретного канала связи является наиболее простой, и возможность аппроксимации канала данной моделью позволяет достаточно быстро получить конечные результаты по характеристикам канала связи. Например, матрица двоичного симметричного канала имеет вид

$$P = \begin{Vmatrix} 1-p & p \\ p & 1-p \end{Vmatrix},$$

где $p$ — вероятность трансформации символа; $1-p$ — вероятность прохождения символа.

Для улучшения свойств двоичного симметричного канала может быть введена зона стирания. Тогда получаем матрицу вида

$$P = \begin{Vmatrix} 1-p-q & p & q \\ p & 1-p-q & q \end{Vmatrix},$$

где $q$ — вероятность стирания символа.

Следует отметить, что при определенных условиях и правильном выборе кода и порога срабатывания решающей схемы

применение канала со стиранием позволяет уменьшить вероятность ошибки в приеме информации. Численные значения вероятностей, характеризующих канал связи, могут быть найдены, если известны конструкция первой решающей схемы и помехи в канале связи. В простейшем случае, когда параметры канала постоянны и помехи могут быть охарактеризованы как стационарные, получаем модель дискретного, в ряде случаев симметричного, канала без памяти.

Таким образом, граф, отображающий дискретный канал, является моделью, или математическим отображением, реального канала связи. На основании этого графа или матрицы вероятностей переходов удается рассчитывать основные характеристики канала, связывающие его с источником информации и позволяющие оценить возможность передачи заданного количества информации по каналу связи. Если все вероятности, входящие в матрицу $P$, кроме стоящих по диагонали, равны нулю, то такой канал является каналом без шума. Основная проблема передачи информации в каналах без шума — это выбор оптимального кода, который по своим свойствам согласуется с источником информации и имеет наименьшую среднюю длину, а в каналах с шумом — выбор кода, который обеспечивает передачу информации с заданной верностью при максимальной скорости передачи. Такая постановка задачи вполне реальна и свойственна для всех практических задач передачи информации в АСУ, хотя в теоретическом плане может встречаться постановка, связанная с обеспечением максимума скорости передачи информации при отсутствии ограничения на верность.

## § 2. Простейшая структурная схема формирования, передачи и выявления сообщений

В АСУ существует разветвленная сеть обмена информацией, в которой весьма ярко представлены такие фазы преобразования информации, как формирование, передача и выявление отдельных информационных потоков.

Основные принципы передачи информации могут быть выявлены на основе простейшей схемы, представленной на рис. 18. Источник информации (ИИ) выдает некоторое сообщение, которое может передаваться механическим или электрическим сигналом. Реальный источник информации в АСУ характеризуется последовательностью сообщений, т. е.

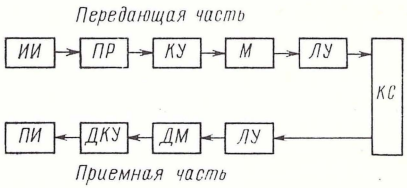


Рис. 18.

может быть указан закон возникновения сообщений во времени, который является в большинстве случаев вероятностным, и при характеристике источника информации обычно отрабатывается некоторая модель источника.

На выходе источника информации возникает совокупность сообщений, которые с переменной периодичностью поступают на преобразователь (Пр), где сообщение преобразуется в электрический сигнал, причем преобразование может быть линейным или нелинейным. На выходе преобразователя один из параметров сигнала зависит от номера передаваемого сообщения. Одним из видов сигнала на выходе преобразователя может быть неизбыточный двоичный код, однозначно отображающий сообщение. В АСУ в качестве преобразователей, осуществляющих подготовку информации, могут быть использованы различные клавишные входные машины, пишущие машинки и другие устройства, позволяющие осуществить преобразование входных сообщений.

Для того чтобы ввести избыточность, необходимую для достоверной передачи сигнала по каналу связи, на передающей стороне предусматривается кодирующее устройство (КУ), с помощью которого формируется избыточный код. В зависимости от выбора могут быть сформированы двоичный, троичный, четверичный коды, код с параллельной или последовательной выдачей элементов во времени, а также код, использующий различные сигнальные признаки. Введение избыточности в код осуществляется именно в кодирующем устройстве, уровень избыточности должен быть согласован с характеристиками канала связи и принятыми алгоритмами передачи информации по каналу.

В настоящее время наиболее широкое применение в задачах АСУ нашли кодирующие устройства, выполненные на регистрах сдвига и предназначенные для формирования циклических кодов различной длины. Использование циклических кодов позволяет осуществить достаточно оперативно и в простой форме кодирование как каждого сообщения, так и нескольких сообщений последовательно единым кодом. Повышение длины используемого кода способствует оптимальному введению избыточности в код и уменьшает среднюю долю избыточных символов среди общего количества символов в коде. Таким образом, на выходе кодирующего устройства получаем совокупность импульсов, отображающих код одного или нескольких сообщений. На нижних ступенях информационной сети АСУ этот код может быть передан непосредственно в канал связи без дополнительного преобразования. Однако при этом реализуется малая скорость передачи информации, и поэтому на верхних ступенях информационной сети АСУ в целях повышения быстродействия системы передача кодов происходит с дополнительным преобразованием в виде модуляции. Для этого служит модуля-

тор (*М*), с помощью которого осуществляется воздействие импульсов кода, как управляющего сигнала, на один из параметров переносчика. С выхода М снимается модулированный сигнал, который через линейное устройство (ЛУ), осуществляющее согласование передающей части с каналом связи, поступает в канал связи (КС).

Следует отметить, что при использовании конкретной линии связи на ней может быть осуществлена реализация одного или более каналов связи, поэтому если линия связи используется для передачи лишь одной информации, то понятие «канал связи» становится тождественным понятию «линия связи». При многоканальной передаче информации в одной линии связи возникает несколько каналов связи. В канале связи на передаваемую информацию воздействует помеха. Для оценки канала связи и действия помех обычно строятся модели непрерывного и дискретного каналов и последствий действия помех в виде модели ошибок.

На нижнем уровне передачи информации в АСУ существуют специально выделенные линии связи, и здесь могут быть использованы модели независимых ошибок, т. е. модели пуассоновского закона распределения ошибок во времени. На верхних ступенях передачи информации вследствие использования высокоскоростной передачи и применения модемов, использования коммутируемых линий связи наблюдается группирование ошибок, и должны быть рассмотрены модели более сложных каналов связи и, соответственно, модели более сложных потоков ошибок, характеризующие различные состояния канала и разные варианты группирования ошибок в канале.

С выхода канала сигнал поступает на ЛУ, согласующее канал связи с приемной частью, демодулируется с помощью демодулятора (ДМ). Выявленные импульсы кода поступают в декодирующее устройство (ДКУ). В ДКУ возможна реализация различных алгоритмов выявления сообщений, в связи с чем может быть осуществлено при той же самой избыточности либо только обнаружение, либо обнаружение и исправление ошибок. Выявленная информация поступает на потребитель информации (ПИ). Следует отметить, что рассмотренная схема формирования, передачи и выявления информации является простейшей. В ней не предусматривается использование дополнительных средств повышения верности передачи.

Использование обратной связи весьма существенно влияет на повышение верности. Следует отметить, что введение обратной связи обычно возможно только в том случае, когда в реальных каналах связи предусматривается дуплексный режим работы, т. е. передача информации в обратном направлении. Современные стандартные устройства уплотнения каналов связи — модемы — позволяют осуществить на одном комплексе аппаратуры высокоскоростную передачу информации в прямом

направлении и низкоскоростную передачу информации в обратном. Иначе говоря, непосредственно в алгоритм функционирования модемов закладывается возможность реализации системы с обратной связью, а поэтому использование систем с обратной связью в АСУ является весьма актуальным и позволяет реализовать требуемую верность передачи информации.

Рассмотрим структурную схему системы с обратной связью. Принципиально все системы с обратной связью могут быть разделены на две большие группы: системы с решающей обратной связью (РОС) и системы с информационной обратной связью (ИОС). Различие между этими системами заключается в месте принятия решения о необходимости повторной передачи: если решение принимается на приемной стороне, то имеем систему с РОС, если на передающей, то возникает система с ИОС.

Система с автоматическим запросом представляет собой одну из элементарных и наиболее распространенных модификаций систем с обратной связью. В настоящее время она используется также для передачи цифровой информации, данных между вычислительными комплексами и т. п. Система с запросом может быть выполнена либо с покодовой, либо с поэлементной проверкой сообщения.

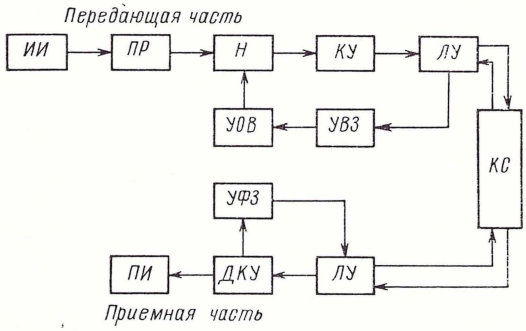


Рис. 19.

Структурная схема системы с запросом с покодовой проверкой сообщения представлена на рис. 19. Источник информации (ИИ) выдает сообщение на преобразователь (Пр). Накопитель (Н) согласует поток сообщений от ИИ с возможностями передачи по линии связи. Кодирование сообщений в некотором избыточном коде осуществляет кодирующее устройство (КУ), которое выдает код сообщения в ЛУ, обеспечивающее согласование передающей части системы с линией связи. На приемной стороне код сообщения через ЛУ поступает на ДКУ, которое может быть настроено на обнаружение или на исправление ошибок. При наличии обнаруживаемой ошибки ДКУ выдает сигнал на устройство формирования запроса (УФЗ), которое

посылает код или импульс запроса в линию связи. На передающей стороне запрос выявляется устройством выявления запроса (УВЗ), которое выдает сигнал на устройство однократной выдачи сообщения (УОВ). Последнее обеспечивает однократное повторение предыдущего сообщения из накопителя. При отсутствии ограничений на число передач данное сообщение повторяется, пока не прекращается запрос с приемной стороны. Таким образом, при покодовой проверке сигнал запроса посылается после обнаружения ошибки в коде сообщения.

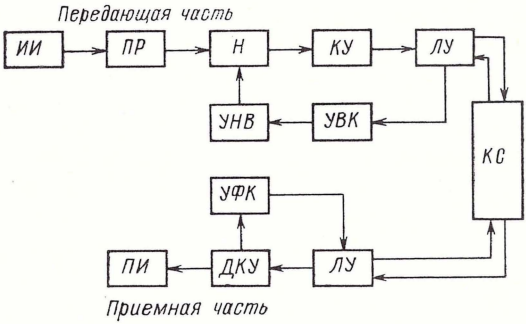


Рис. 20.

Другой простейшей разновидностью систем с РОС является система с квитированием. Она может быть выполнена с покодовой или поэлементной проверкой сообщения на приемной стороне. Структурная схема системы с квитированием с покодовой проверкой сообщения представлена на рис. 20. Декодирующее устройство воздействует на устройство формирования квитанции (УФК) в случае, если принятая комбинация попала в одну из разрешенных. На передающей стороне код или импульс квитанции фиксируется устройством выявления квитанции (УВК), которое воздействует на устройство неограниченной выдачи (УНВ), прекращая дальнейшее повторение данного сообщения. Таким образом, квитанция передается по обратному каналу как в случае правильного приема, так и при трансформации сообщения.

В настоящее время в системе с РОС широко используется передача по обратному каналу сигнала запроса и сигнала квитирования. Используя рассмотренные выше простейшие структурные схемы, можно составить общую схему, объединяющую оба варианта.

Простейшим вариантом системы с ИОС является система со сравнением, где обратной связью охватываются прямой канал передачи информации, линейное и декодирующее устройства. На рис. 21 представлена структурная схема системы

с покодовым сравнением без передачи результатов сравнения по прямому каналу. Устройство неограниченной выдачи сообщений (УНВ) обеспечивает многократную передачу одного и того же сообщения из накопителя (Н). Каждое сообщение кодируется кодирующим устройством (КУ) и через линейное устройство (ЛУ) выдается в линию связи. На приемной стороне прямого канала декодирующее устройство (ДКУ) выявляет принятое сообщение. Это сообщение повторно кодируется кодирующим устройством (КУ) и посылается по каналу обратной связи. На передающей стороне с помощью схемы сравнения (СС) устанавливается соответствие переданного сообщения с принятым из обратного канала. При отсутствии искажений схема сравнения тормозит устройство неограниченной выдачи. В линии связи прекращается передача информации. Зафиксированное ранее сообщение в запоминающем устройстве (ЗУ) выдается на исполнительное устройство. При отсутствии временного интервала, что имеет место в случае многократной передачи одного и того же сообщения, запоминающее устройство в начале каждого цикла стирает предыдущее и фиксирует следующее сообщение.

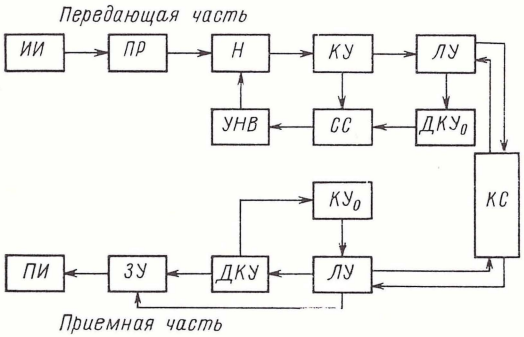


Рис. 21.

Возможны различные варианты реализации сравнения. В зависимости от этого меняется верность передаваемой информации и вероятностные показатели функционирования системы. Существенным достоинством системы с ИОС является малая вероятность трансформации информации в системе ввиду малой вероятности трансформации однонаправленного типа в прямом и обратном каналах. На основе рассмотренных вариантов систем могут быть построены системы с комбинированной обратной связью, с блокировкой, с адресным повторением и т. д. Компонуя элементарные простейшие структурные схемы передачи информации в единую совокупность, получаем информационную сеть, представляющую собой соединение набора источников информации, потребителей и различных узлов объединения информационных потоков.

## § 3. Скорость передачи информации и пропускная способность канала связи

Основным назначением любого канала связи является передача информации с заданным качеством. Используя введенное ранее понятие количественной меры информации на основе энтропии, нетрудно найти количество информации, передаваемое по каналу связи. Если дискретный канал отображается некоторым графом переходов и на его входе имеется совокупность символов $x \in X$, а на выходе выявляется совокупность символов $y \in Y$, то неопределенность этого выявления можно оценить безусловной энтропией $H(Y)$. Неопределенность, вызванную действием помех в канале связи, соответственно можно оценить условной энтропией $H(Y/X)$, и эта неопределенность характеризует численно ту долю информации, которая теряется в канале связи под действием помех.

Поскольку ввиду прихода информации из канала снимается неопределенность на выходе за исключением той доли неопределенности, которую вносит помеха, то количество информации, передаваемое по каналу связи [22]

$$I = H(Y) - H(Y/X),$$

$H(Y)$ зависит от статистики появления символов на выходе канала. Если имеет место взаимно-однозначное кодирование информации, то $H(Y) = H(X)$, $H(Y/X)$ зависит также от свойств источника, но здесь весьма существенными оказываются характеристики канала связи и правила кодирования.

Целесообразно для любого канала связи обеспечить максимум количества передаваемой информации, а поэтому одной из характеристик канала связи являются соотношения, связанные с предельными значениями количества передаваемой информации.

*Пропускная способность дискретного канала связи без шума.* При отсутствии шума $H(Y/X) = 0$, тогда

$$I = H(Y) = -\sum_{j=1}^{K_x} \sum_{i=1}^{K_y} p(x_j) \log_2 [p(x_i)\, p(y_i/x_j)] =$$
$$= -\sum_{i=1}^{K_y} p(y_i) \log_2 p(y_i).$$

Верхний предел количества информации, которую может передать канал связи, называется пропускной способностью канала связи:

$$C = I_{\max} = H_{\max}(Y).$$

Если считать, что символы на выходе канала связи равновероятны, то $p(y_i) = 1/K_y$, тогда пропускная способность канала без шума

$$C = H_{\max}(Y) = -K_y (1/K_y) \log_2 (1/K_y) = \log_2 K_y.$$

На рис. 22 представлена зависимость пропускной способности канала связи от основания кода: чем больше основание кода, тем более информационным является символ кода, однако усложняется схемная реализация.

*Пропускная способность дискретного канала связи с шумом.* Пусть имеется канал без стирания, т. е. $K_x = K_y = K$, тогда

$$I = H(Y) - H(Y/X) = -\sum_{j=1}^{K}\sum_{i=1}^{K} p(x_j)\,p(y_i/x_j) \times$$
$$\times \log_2 [p(x_j)\,p(y_i/x_j)] + \sum_{j=1}^{K}\sum_{i=1}^{K} p(x_j) \times$$
$$\times p(y_i/x_j) \log_2 p(y_i/x_j).$$

Используя полученную формулу, рассмотрим $K$-ичный симметричный канал, граф которого изображен на рис. 23. Обозначим через $p$ вероятность трансформации символа, тогда $\sum_{i=1}^{k} p(y_i/x_j) = p$. Отсюда $p(y_i/x_j) = p/(K-1)$, $i \neq j$, — вероятность трансформации символа $x_j$ в символ $y_i$; $p(y_i/x_j) = 1-p$, $i = j$, — вероятность прохождения символа. Пропускная способность этого канала связи

$$C = I_{\max} = H_{\max}(Y) - H_{\min}(Y/X).$$

Для получения $H_{\max}(Y)$ принимаем, что символы на выходе равновероятны, т. е. $p(y_i) = 1/K$, тогда

$$H_{\max}(Y) = -\sum_{i=1}^{K} p(y_i) \log_2 p(y_i) = \log_2 K.$$

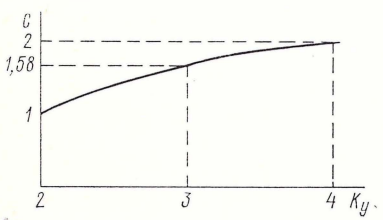


Рис. 22.

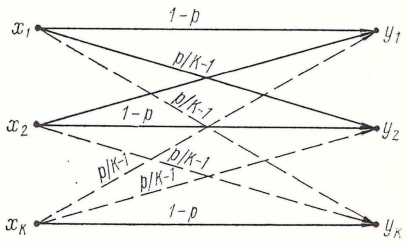


Рис. 23.

Определим $H_{\min}(Y/X)$, для этого найдем

$$H(Y/X) = \sum_{j=1}^{K}\sum_{i=1}^{K} p(x_j)\,p(y_i/x_j) \log_2 p(y_i/x_j).$$

Если символы на входе равновероятны, т. е. $p(x_j) = 1/K$, то

$$H(Y/X) = -K\left(\frac{1}{K}\right)\left[(1-p)\log_2(1-p) + \right.$$
$$\left. + (K-1)\frac{p}{K-1}\log_2\frac{p}{K-1}\right] =$$
$$= (1-p)\log_2(1-p) - p\log_2\frac{p}{K-1}.$$

Ниже будет показано, что при соответствующем выборе порога срабатывания можно получить минимум условной энтропии. Окончательно находим, что

$$C = \log_2 K + (1-p)\log_2(1-p) + p\log_2\frac{p}{K-1}.$$

Физически эта величина отображает максимально возможное количество информации, которое может пропустить канал связи в одном символе кода, если в канале действует помеха и вероятность искажения символа равна $p$.

В частном случае двоичного симметричного канала, для которого $K=2$, пропускная способность

$$C = 1 + (1-p) \times \times \log_2(1-p) + p\log_2 p.$$

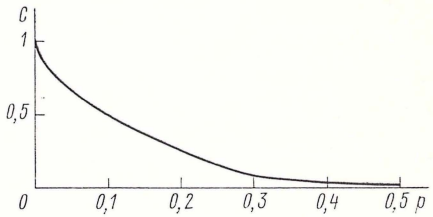


Рис .24.

На рис. 24 представлена зависимость пропускной способности канала связи от вероятности искажения символа. Реальные системы работают при $p = 10^{-3} \div 10^{-5}$.

Свяжем пропускную способность с физическими характеристиками канала $T_к$, $F_к$, $H_к$. Для этого перейдем от $C$ $\left[\frac{\text{дв. единиц}}{\text{символ}}\right]$ к $C_0$ $\left[\frac{\text{дв. единиц}}{\text{с}}\right]$, $C_0 = Cv$, где $v$ — число символов, передаваемых по каналу связи в одну секунду. Тогда $v = 1/\tau = F_к$, где $\tau$—длительность символа. Можно записать, что $C_0 = F_к C$ $\left[\frac{\text{дв. единиц}}{\text{с}}\right]$. Будем считать, что помеха и сигнал взаимно независимы и разделимы. Тогда количество информации, поступающей в приемник:

$$I_0 = I_с + I_п,$$

где $I_с$ — количество сведений сигнала, $I_п$ — количество сведений помехи.

К. Шенноном показано [28], что в общем случае при ограниченной мощности сигнала количество сведений, которое несет сигнал с длительностью $T$, шириной полосы $F$ и мощностью $P_с$, определяется как

$$I = TF\log_2 AP,$$

где $A$ — коэффициент пропорциональности. Для увеличения количества сведений при отсутствии помех можно в равной

мере увеличивать длительность, ширину полосы или мощность сигнала. С учетом воздействия помехи получаем

$$I_c = I_0 - I_п = TF\log_2 A\,(P_c + P_п) - TF\log_2 AP_п = \\ = TF\log_2 (1 + P_c/P_п).$$

Тогда если $P_c < P_п$, то передача возможна, однако количество сведений резко уменьшается. Для компенсации необходимо увеличивать длительность либо ширину полосы сигнала:

$$I_{max} = C = T_к F_к \log_2 (1 + P_c/P_п).$$

Рассмотрим, как зависит пропускная способность канала $C$ от основания кода $K$, исходя из физических свойств.

Если $E_c$ — энергия сигнала, то $E_c = P_c T = P_c T_к$ при $T = T_к$; $P_п/F_к = P_{п0}$, где $P_{п0}$ — удельная мощность помехи, т. е. мощность, приходящаяся на единицу полосы. Можно записать, что

$$P_c/P_п = E_c/T_к F_к P_{п_0}.$$

Если $T_к = T$ и $F_к = F$, то $T_к F_к = TF = b$, где $b$ — база сигнала. Окончательно получаем

$$C = b\log_2 (1 + E_c/b\,P_{п_0}).$$

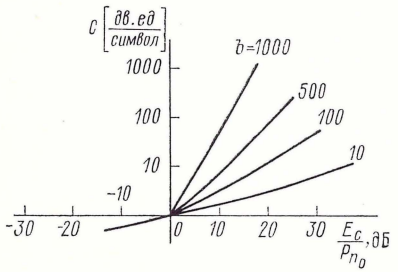


Рис. 25.

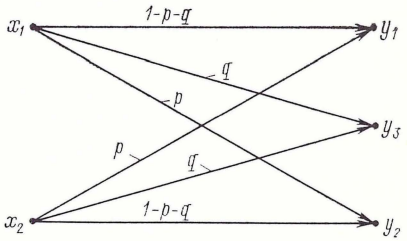


Рис. 26.

На рис. 25 представлены зависимости $C$ от $E_c/P_{п0}$. Из графика следует:

1) идеальная система передачи для диапазона $(0 \div 10)$ дБ должна нести более одной двоичной единицы информации в символе, т. е. система должна быть недвоичной;

2) идеальная двоичная система отображается горизонтальной прямой $C = 1$. Поскольку в канале с шумом $C < 1$, то необходимо вводить избыточность;

3) при значительных превышениях сигнала над помехой и использовании недвоичных кодов целесообразно увеличивать базу сигналов $b$, тогда $C \gg 1$.

Рассмотрим дискретный канал связи со стиранием. Будем считать, что введение зоны стирания не устраняет трансформации символа, а лишь уменьшает ее вероятность. Пусть $q$ — вероятность попадания сигнала в зону стирания. Получаем граф канала, представленный на рис. 26 для $K=2$.

Определим пропускную способность при равновероятности символов на входе:

$$C = H_{\max}(Y) - H_{\min}(Y/X),\ H(Y) = -\sum_{i=1}^{3} p(y_i)\log_2 p(y_i);$$

$$p(y_1) = p(y_2) = p(x_1)\,p(y_1/x_1) + p(x_2)\,p(y_1/x_2) =$$
$$= (1-p-q)/2 + p/2 = (1-q)/2;$$

$$p(y_3) = p(x_1)\,p(y_3/x_1) + p(x_2)\,p(y_3/x_2) = q/2 + q/2 = q;$$

$$H(Y) = [-2(1-q/2)/2]\log_2[(1-q)/2] - q\log_2 q =$$
$$= (1-q)[1-\log_2(1-q)] - q\log_2 q.$$

Найдем условную энтропию

$$H(Y/X) = -\sum_{j=1}^{2}\sum_{i=1}^{3} p(x_j)\,p(y_i/x_j)\log_2 p(y_i/x_j) =$$
$$= -2\,\frac{(1-p-q)\log_2(1-p-q) + p\log_2 p + q\log_2 q}{2} =$$
$$= (1-p-q)\log_2(1-p-q) + p\log_2(p) + q\log_2 q.$$

Окончательно получаем

$$C = (1-q)[1-\log_2(1-q)] +$$
$$+ (1-p-q)\log_2(1-p-q) + p\log_2 p.$$

Для проверки подставим $q=0$ и получим пропускную способность для канала без стирания.

Введение зоны стирания снижает пропускную способность канала связи, если в канале отсутствуют помехи ($C=1-q$ при $p=0$). Целесообразно организовать канал со стиранием при наличии помех, если путем выбора ширины зоны стирания удается обеспечить $p \ll q$.

## § 4. Методы приближения скорости передачи информации к пропускной способности дискретного канала связи

Возможны два метода воздействия на пропускную способность, поскольку она складывается из составляющих $H(Y)$ и $H(Y/X)$.

Рассмотрим, как получить $H_{\max}(Y)$, обеспечивая равновероятность выходных символов, если символы на выходе

не равновероятны [7]. Пусть имеется двоичный несимметричный канал (рис. 27), $H_{\max}(Y) = 1$, если $p(y_1) = p(y_2)$, т. е.

$$p(x_1)\,p(y_1/x_1) + p(x_2)\,p(y_1/x_2) = p(x_1)\,p(y_2/x_1) + p(x_2)\,p(y_2/x_2),$$

или $p(x_1)(1-p_1) + p(x_2)\,p_2 = p(x_1)\,p_1 + p(x_2)(1-p_2)$.

Найдем вероятность трансформации символа

$$p_1 = p(x_2)\,p_2/p(x_1) + 1/2 - p(x_2)/2p(x_1).$$

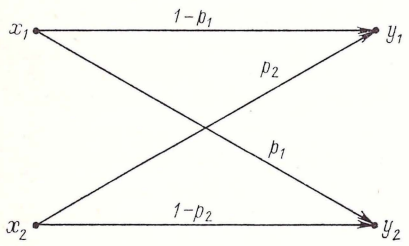


Рис. 27.

Из этого выражения следует:

1) если входные символы равновероятны, т. е. $p(x_1) = p(x_2)$, то $p_1 = p_2$, т. е. максимум энтропии обеспечивается лишь в симметричном канале. Отсюда вытекает, что не всегда равновероятность входных символов является полезной;

2) если входные символы неравновероятны, то пороги срабатывания приемной схемы надо подобрать так, чтобы выполнилось полученное выше условие.

Пример 1. Пусть $p(x_1) = p(x_2)$ и $p_1 > p_2$. Тогда порог срабатывания нужно увеличить для обеспечения условия $p_1 = p_2$. Численное значение порога можно установить, если известны закон распределения помехи и закон распределения помехи и сигнала.

Теперь попытаемся обеспечить $H_{\min}(Y/X)$:

$$H(Y/X) = -\sum_{j=1}^{K}\sum_{i=1}^{K} p(x_j)\,p(y_i/x_j)\log_2 p(y_i/x_j) =$$

$$= -p(x_1)\,[(1-p_1)\log_2(1-p_1) + p_1\log_2 p_1] - p(x_2)\times$$

$$\times\,[(1-p_2)\log_2(1-p_2) + p_2\log p_2] = p(x_1)\,H_1 + p(x_2)\,H_2,$$

где $H_1$ — энтропия двух событий, из которых одно возникает с вероятностью $p_1$, а другое с вероятностью $1-p_1$; $H_2$ — энтропия двух событий, из которых одно возникает с вероятностью $p_2$, а другое с вероятностью $1-p_2$. На рис. 28 представлена зависимость $H_1$, $H_2$ от $p_1$, $p_2$. Будем считать, что символы на входе равновероятны, тогда

$$H(Y/X) = (H_1 + H_2)/2.$$

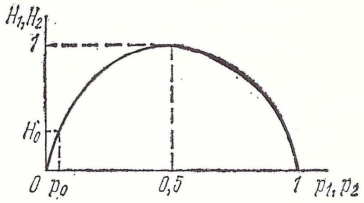


Рис. 28.

Поскольку условная энтропия есть количество информации, которое вносит помеха,

то целесообразно иметь рабочую точку $p_0$, близкую к нулю.

Определим, как зависит условная энтропия от вероятностей $p_1$ и $p_2$. Для этого разложим функции $H_1$ и $H_2$ в окрестности точки $p_0$ в ряд Тейлора, ограничиваясь двумя первыми членами:

$$H_1 = H_0 + \Delta p_1 \frac{dH_1}{dp}\Big|_{p=p_0}, \quad H_2 = H_0 + \Delta p_2 \frac{dH_2}{dp}\Big|_{p=p_0}.$$

Отсюда $H(Y/X) = H_0 + \frac{\Delta p_1 + \Delta p_2}{2} \frac{dH}{dp}\Big|_{p=p_0}$. Исследуем это выражение на минимум. Так как $dH/dp$ положительна в точке $p_0$, то желательно иметь минимум приращения $\Delta p_1 + \Delta p_2$. Приращения возникают, если меняется порог срабатывания приемной схемы. Найдем значение порога, соответствующее минимуму условной энтропии. Передача информации обычно осуществляется по каналу с шумом. Распределение помехи с шумом в отсутствие сигнала обозначим $W(x)$, распределение помехи и сигнала с амплитудой $a$ обозначим $W_a(x)$. Различают два случая распределений: симметричные и несимметричные. Симметричное распределение представлено на рис. 29:

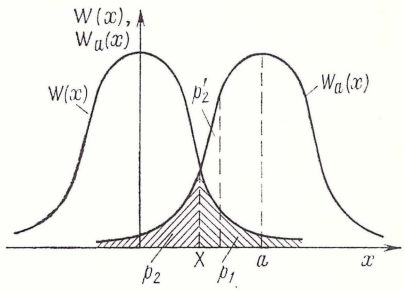

Рис. 29.

$$p_1 = p(y_2/x_1) = p(1/0), \quad p_2 = p(y_1/x_2) = p(0/1),$$

$$p_1 = \int_X^\infty W(x)\,dx, \quad p_2 = \int_{-\infty}^X W_a(x)\,dx.$$

Сместим порог срабатывания вправо, т. е. увеличим $X$, тогда $p_1' < p_1$, т. е. $\Delta p_1 < 0$, $p_2' > p_2$, т. е. $\Delta p_2 > 0$. Так как $|\Delta p_2| > |\Delta p_1|$, то $\Delta p = \Delta p_1 + \Delta p_2 > 0$, т. е. при сдвиге порога вправо условная энтропия увеличилась.

Сместим порог влево, т. е. уменьшим $X$, тогда $p_1'' > p_1$, $\Delta p_1 > 0$; $p_2'' < p_2$, $\Delta p_2 < 0$. Так как $|\Delta p_1| > |\Delta p_2|$, то $\Delta p = \Delta p_1 + \Delta p_2 > 0$, т. е. при сдвиге порога влево условная энтропия увеличилась. Минимум условной энтропии наблюдается, если порог срабатывания выбран в точке пересечения кривых $W(x)$ и $W_a(x)$.

Пример 2. Пусть передача осуществляется импульсами постоянного тока с амплитудой $a$ (рис. 30). В канале связи (КС) действует нормальный шум. Пороговая схема (ПС) обеспечивает выявление элементов кода:

$$W(x) = (1/\sqrt{2\pi\sigma^2})\exp(-x^2/2\sigma^2); \quad W_a(x) = (1/\sqrt{2\pi\sigma^2})\exp[-(x-a)^2/2\sigma^2],$$

где $\sigma$ — эффективное напряжение помехи. Порог срабатывания нужно выбирать равным половине амплитуды сигнала на выходе приемника, т. е. $X = a/2$.

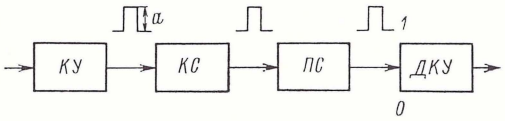


Рис. 30.

Несимметричные распределения имеют место в условиях действия нормального шума, если на входе приемной схемы устанавливаются фильтры, т. е. передача элементов кода осуществляется импульсами переменного тока заданной частоты. Распределения могут быть несимметричными и в случае, когда закон распределения помехи в канале отличается от нормального. Можно показать (рис. 31), что порог срабатывания также необходимо выбрать в точке пересечения этих распределений. Порог перемещается с изменением амплитуды сигнала по нелинейному закону. При каждом новом значении амплитуды приходится решать уравнение для оценки порога срабатывания.

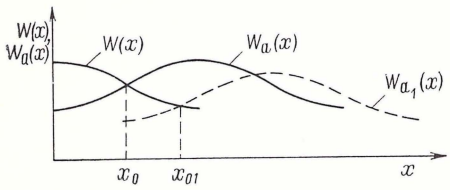


Рис. 31.

Пример 3. Пусть передача осуществляется по каналу с нормальным шумом импульсами переменного тока и на входе приемной схемы установлен фильтр (рис. 32): $М$ — модулятор; $Ф$ — фильтр, разделяющий частоты; $Д$ — детектор; $Н$ — накопитель, устраняющий пульсацию; $ПС$ — пороговая схема, формирующая импульс прямоугольной формы.

Из теории случайных процессов известно, что при наложении нормального шума на синусоидальные колебания вероятность появления заданной амплитуды подчиняется распределению Релея на выходе фильтра [23]:

$$W(x) = (x/\sigma^2)\exp(-x^2/2\sigma^2).$$

С учетом сигнала

$$W_a(x) = (x/\sigma^2) \times \exp[(-x^2-a^2)/2\sigma^2]\, I_0(ax/\sigma^2),$$

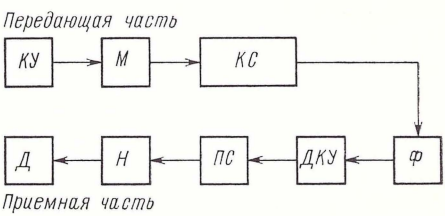


Рис. 32.

где $I_0$ — функция Бесселя нулевого порядка, $a$ — амплитуда колебаний полезного сигнала.

Перейдем к нормированным величинам:

$$W(x)\,dx = W(y)\,dy,$$

где $y$ — нормированная величина; $y = x/\sigma$, тогда $dy = dx/\sigma$, $dx = \sigma dy$, $W(x)\,\sigma dy = W(y)\,dy$, $W(y) = W(x)\,\sigma$, $W(y) = y\exp(-y^2/2)$,

$$W_a(y) = y\,I_0(ay/\sigma)\exp[-(y^2 + a^2/\sigma^2)/2].$$

Введем величину $\delta = P_c/P_п$, где $P_c$ и $P_п$ — мощности сигнала и помехи соответственно, $\delta = (a/\sqrt{2})^2/\sigma^2$. Так как $a/\sigma = \sqrt{2\delta}$, то

$$W_a(y) = yI_0(y\sqrt{2\delta})\exp[(-y^2/2) - \delta].$$

На рис. 33 представлены зависимости $W(y)$, $W_a(y)$ от $y$; $W_a(y) = W(y)$ при $y = Y$, тогда

$$Y\exp\left[\left(-\frac{Y^2}{2}\right) - \delta\right] I_0(Y\sqrt{2\delta}) = Y\exp\left(\frac{Y^2}{2}\right),$$

$$I_0(Y\sqrt{2\delta})\exp(-\delta) = 1.$$

При увеличении $\delta$ значение $Y$ возрастает, вероятность ошибки уменьшается.

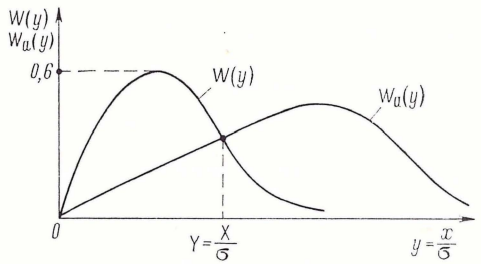


Рис. 33.

## § 5. Скорость передачи информации и пропускная способность непрерывного канала связи

Значительная часть информации в АСУ имеет непрерывный характер. При передаче этой информации возникают сообщения с непрерывным распределением элементов. Определим количество информации, передаваемое через канал связи. Для этого воспользуемся полученным в § 3 гл. III соотношением $I = H(Y) - H(Y/X)$, где $H(Y)$ — безусловная энтропия на выходе канала связи, зависящая от плотности распределения вероятности $W(y)$ непрерывного сообщения, принятого с выхода канала; $H(Y/X)$ — условная энтропия, отображающая потери информации при передаче через канал связи.

В канале связи без шума $H(Y/X)=0$, количество передаваемой информации

$$I=H(Y)=-\int_{-\infty}^{\infty} W(y)\log_2[W(y)\Delta y]\,dy,$$

где $\Delta y$ — интервал между элементами сообщения. Как показано выше, наибольшее количество информации соответствует оптимальному распределению $W_{\text{опт}}(y)=\left(1/\sqrt{2\pi\sigma^2}\right)\exp(-y^2/2\sigma^2)$. В канале без шума $W(y)=W(x)$, т. е. сохраняется распределение значений функции, передаваемых в канал связи. Так как предельное значение количества передаваемой информации соответствует пропускной способности канала связи, то при $W(y)=W_{\text{опт}}(y)$ получаем

$$C=I_{\max}=H_{\max}(Y)=\log_2\left(\sigma\sqrt{2\pi e}/\Delta x\right),$$

где $\Delta x$ определяется с учетом требуемой точности воспроизведения непрерывной функции.

Зачастую пропускную способность непрерывного канала без шума находят из условия максимума приведенной энтропии. Тогда $\Delta x$ исключается из рассмотрения и пропускная способность $C=\log_2\left(\sigma\sqrt{2\pi e}\right)$.

В целях приближения скорости передачи информации к пропускной способности канала необходимо так осуществлять ввод информации в канал связи, чтобы значения непрерывной функции подчинялись нормальному центрированному распределению.

В канале связи с шумом $H(Y/X)\neq 0$. Количество передаваемой информации

$$I=H(Y)-H(Y/X)=-\int_{-\infty}^{\infty} W(y)\log_2[W(y)\Delta y]\,dy+$$

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x)\,W(y/x)\log_2[W(y/x)\Delta y]\,dxdy.$$

Наличие помехи снижает количество информации, передаваемой по каналу связи. Представим энтропию в виде

$$H(Y)=-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x)\,W(y/x)\log_2[W(y)\Delta y]\,dxdy.$$

Обозначим $W(x)\,W(y/x)=W(y,x)$. Тогда

$$I=-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(y,x)\log_2[W(y)\Delta y]\,dxdy+$$

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(y,x)\log_2\left[\frac{W(y,x)\Delta y}{W(x)}\right]dxdy=$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(y,x)\log_2\left[\frac{W(y,x)}{W(x)\,W(y)}\right]dxdy.$$

Отметим, что в предельном случае, когда сообщения на входе и на выходе канала статистически независимы (помеха подавляет сигнал в канале), $W(y, x) = W(x)\, W(y)$ и, как следствие, $I = 0$. Пропускная способность непрерывного канала с шумом может быть определена как предельное значение количества передаваемой информации (см. [28]):

$$C = I_{\max} = \max \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(y, x) \log_2 \left[ \frac{W(y, x)}{W(x)\, W(y)} \right] dx dy \right\}.$$

Максимум ищется по всевозможным распределениям значений непрерывной функции на входе канала связи.

# *Глава IV*

## ДИСКРЕТНАЯ ИНФОРМАЦИЯ В АСУ

### § 1. Кодирование дискретной информации в каналах связи без шума

Канал связи без шума — это канал, в котором помехи либо отсутствуют, либо ими можно пренебречь. Основной проблемой кодирования в канале связи без шума является обеспечение скорости передачи информации, сколь угодно близкой к его пропускной способности.

К. Шенноном доказано [28], что для источника сообщений с энтропией $H(X)$ можно так построить код, чтобы передавать сообщения по каналу связи со средней скоростью $H(X) = C - \varepsilon$, где $C$ — пропускная способность канала без шума; $\varepsilon$ — сколь угодно малое число. Доказательство теоремы может быть выполнено с учетом такого способа кодирования, при котором наиболее вероятные сообщения отображаются кодами наименьшей длины, и наоборот. Тогда сообщению $x_{0j}$, возникающему с вероятностью $p(x_{0j})$, соответствует код длиной $n_j$. Будем считать, что $p(x_{01}) \geqslant p(x_{02}) \geqslant \ldots \geqslant p(x_{0j}) \geqslant \geqslant \ldots \geqslant p(x_{0м})$, и построим двоичные коды для каждого сообщения как разложение их вероятностей в виде двоичных чисел. Тогда для сообщения $x_{0j}$ получим код длиной $n_j$, причем из условия разложения следует, что

$$\log_2 [1/p(x_{0j})] \leqslant n_j < 1 + \log_2 [1/p(x_{0j})], \text{ или}$$

$$1/2^{n_j} \leqslant p(x_{0j}) < 1/2^{n_j - 1}.$$

Если сообщение имеет длину $n$, то количество информации, содержащееся в одном символе исходного сообщения:

$$I = (1/n) \sum_{j=1}^{M} n_j p(x_{0j}),$$

где $M$ — число сообщений. Таким образом,

$$\frac{1}{n} \sum_{j=1}^{M} \log_2 \left[ \frac{1}{p(x_{0j})} \right] p(x_{0j}) \leqslant \frac{1}{n} \sum_{j=1}^{M} n_j p(x_{0j}) <$$

$$< \frac{1}{n} \sum_{j=1}^{M} \left\{ 1 + \log_2 \left[ \frac{1}{p(x_{0j})} \right] \right\} p(x_{0j}).$$

С увеличением $n$ (ростом числа сообщений) все члены неравенства стремятся к энтропии источника $H(X)$. Заметим, что $\sum_{j=1}^{M} n_j p(x_{0j}) = n_{ср}$, где $n_{ср}$ — средняя длина кода. Скорость передачи всего множества сообщений $R = H(X_0)/n_{ср}$, где $H(X_0)$ — энтропия сообщения.

Пусть имеется набор сообщений $x_{01}, \ldots x_{0м}$ (рис. 34), $\sum_{j=1}^{M} p_{0j} = 1$. Требуется так построить код, чтобы $n_{ср} \approx n_{опт}$, где $n_{опт}$ — оптимальная (минимально возможная) длина кода:

$$n_{опт} = H(X_0)/H_{max}(X),$$

где $H_{max}(X)$ — максимальное количество информации, которое может передавать один элемент кода.

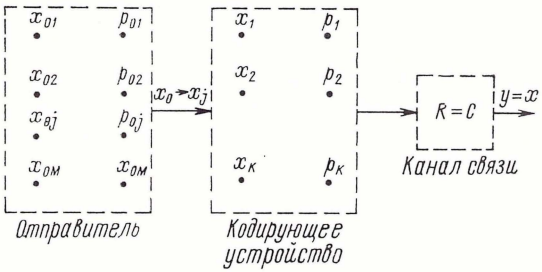


Рис. 34.

Рассмотрим передачу равновероятных сообщений. Попытаемся построить код, исходя из условий: а) символы кода должны быть равновероятны, б) появление последующих символов кода не должно зависеть от предыдущих.

Пример 1. Пусть требуется передать восемь равновероятных сообщений по каналу без шума:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Сообщение . . . . . | $x_{01}$ | $x_{02}$ | $x_{03}$ | $x_{04}$ | $x_{05}$ | $x_{06}$ | $x_{07}$ | $x_{08}$ |
| Вероятность появления . . | 1/8 | 1/8 | 1/8 | 1/8 | 1/8 | 1/8 | 1/8 | 1/8 |
| Код . . . . . | 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 |

Воспользуемся двоичным кодом. Делим сообщения на две равновероятные группы. Присваиваем первой группе символ 0, а второй 1, полученные группы разбиваем на подгруппы, присваивая первой подгруппе символ 0, а второй 1 и т. д. Определяем среднее число нулей, которое имеется во всем ансамбле сообщений:

$$n_0 = \sum_{j=1}^{M} p(x_{0j})\, n_{0j} = (1/8)\, 8 \cdot 12 = 12,$$

где $n_{0j}$ — число нулей в коде $x_{0j}$=го сообщения. Общее число элементов в кодах всех сообщений $nM = 24$, вероятность появления нуля $p(0) = 1/2$, вероятность появления единицы $p(1) = 1/2$. Найдем энтропию источника:

$$H(X) = -\sum_{j=1}^{K} p(x_j) \log_2 p(x_j) = \log_2 K = 1,$$

где $p(x_j)$ — вероятность появления $x_j$-го символа кода, $K$ — основание кода. Так как сообщения равновероятны, то

$$n_{\text{опт}} = H(X_0)/H_{\max}(X) = \log_2 M/\log_2 K = 3.$$

Рассмотрим в общем случае передачу равновероятных сообщений. Пусть имеется $M$ сообщений, которые нужно передавать по каналу без шума, т. е. $p(x_{0j}) = 1/M$. Воспользуемся неизбыточным кодом, который включает всевозможные комбинации символов, имеющих $K$ значений, тогда $M = K^n$. Оптимальная длина кода $n_{\text{опт}} = \log_2 M/\log_2 K$. Длина получаемого кода

$$n = H(X_0)/\log_2 K = \log_2 M/\log_2 K = n_{\text{опт}}.$$

Пример 2. Пусть необходимо передать девять сообщений троичным кодом, т. е. $M = 9$, $K = 3$. Тогда $n = \log_K M = 2$, $n_{\text{опт}} = \log_2 M/\log_2 K = 2\log_2 3/\log_2 3 = 2$.

Рассмотрим кодирование неравновероятных сообщений. В данном случае необходимо получить минимальную среднюю длину кода, т. е. необходимо сообщениям с большей вероятностью появления присваивать меньшую длину кода, и наоборот.

Пример 3. Пусть требуется передать восемь неравновероятных сообщений.

| Сообщение | $x_{01}$ | $x_{02}$ | $x_{04}$ | $x_{04}$ | $x_{05}$ | $x_{06}$ | $x_{07}$ | $x_{08}$ |
|---|---|---|---|---|---|---|---|---|
| Вероятность появления | 1/4 | 1/4 | 1/8 | 1/8 | 1/16 | 1/16 | 1/16 | 1/16 |
| Код | 000 | 01 | 100 | 101 | 1100 | 1101 | 1110 | 1111 |

Делим всю совокупность неравновероятных сообщений на две равновероятные группы и первой присваиваем символ 0, второй 1, затем деление продолжаем до получения в каждой группе одного сообщения. Проверим код на оптимальность:

1) путем сравнения $n_{\text{ср}}$ и $n_{\text{опт}}$:

$$n_{\text{опт}} = H(X_0)/H_{\max}(X) = \left[-\sum_{j=1}^{M} p(x_{0j}) \log_2 p(x_{0j})\right]/\log_2 K = 11/4,$$

$$n_{\text{ср}} = \sum_{j=1}^{M} n_j p(x_{0j}) = 11/4, \quad \text{т. е. } n_{\text{опт}} = n_{\text{ср}};$$

2) путем оценки вероятности появления символов 0 и 1:

$$n_0 = \sum_{j=1}^{M} p(x_{0j})\, n_{0j} = 11/8, \quad p(0) = n_0/n_{\text{ср}} = 1/2, \quad p(1) = 1/2,$$

символы равновероятны, т. е. код оптимальный,

## § 2 Методы построения оптимальных кодов

Построение оптимального кода, т. е. получение условия $n_{\text{ср}} = n_{\text{опт}}$, возможно лишь для отдельных частных распределений вероятностей появления сообщений. Пусть имеется неко-

торый граф, узлы которого отображают отдельные кодовые комбинации. Считаем, что кодовые комбинации располагаются в концевых узлах, а из каждого промежуточного узла отходит $K$ ветвей. Такой граф получил название *кодового дерева*. На рис. 35 дан пример построения кодового дерева для $K=2$. Здесь ни одна комбинация оптимального кода не лежит в промежуточном узле. Поскольку ни одна из этих комбинаций не является началом последующей, то на приемной стороне комбинации кода легко разделяются, и оптимальный код может передаваться без разделительных интервалов. В случае равновероятных сообщений число концевых узлов должно быть равно $K^n$. В случае неравновероятных сообщений не удается использовать концевые узлы одного и того же порядка. В соответствии с этим возникает необходимость общей оценки возможного числа используемых узлов при построении оптимальных кодов неравной длины.

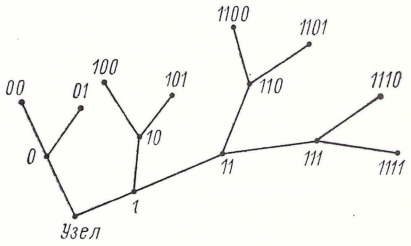

Рис. 35.

**Теорема.** *Необходимым и достаточным условием существования кодовых слов, соответствующих концевым узлам порядка $n_i$, является неравенство Крафта $\sum_{i=1}^{M} K^{-n_i} \leqslant 1$, где $n_i$ — длина кода $i$-го сообщения, $M$ — число передаваемых сообщений* [21].

Доказательство. Необходимо учесть, что при использовании всех возможных концевых узлов получается число комбинаций длиной $n$ не более $K^n$. Но в этом случае все коды имеют одинаковую длину. Поскольку для более вероятных сообщений необходимо использовать коды меньшей длины, то это означает, что вводится ряд концевых узлов порядка $n_i < n$, которые отсекают $K^{n-n_i}$ возможных узлов порядка $n$.

Длины кодовых слов, соответствующих порядку концевых узлов, должны соответствовать условию $\sum_{i=1}^{M} K^{n-n_i} \leqslant K^n$. Делим левую и правую части неравенства на $K^n$ и получаем неравенство, которое требовалось доказать.

Таким образом, условие, сформулированное в данной теореме, является *необходимым* для оценки числа концевых узлов кодового дерева различного порядка. Однако нужно показать, что действительно такое дерево может быть построено и все концевые узлы могут быть использованы, т. е. доказать *достаточность* этого условия. Для этого выделим некоторые концевые узлы порядка $m$, число которых $n_m$. Обозначим через $l$ число концевых узлов порядка $l < m$.

Тогда, исходя из неравенства теоремы, можно записать следующее неравенство:

$$\sum_{i=1}^{l} K^{-n_j} + n_m K^{-m} + \sum_{j=l+n_m+1}^{M} K^{-n_i} \leqslant 1,$$

или $n_m \leqslant K^m - \sum_{i=1}^{l} K^{m-n_i} - \sum_{i=l+n_m+1}^{M} K^{m-n_i}$.

Если учесть, что число возможных для использования концевых узлов порядка $m$ определяется величиной $K^m - \sum_{i=1}^{l} K^{m-n_i}$, где второй член учитывает узлы, исключаемые за счет наличия концевых узлов порядка $l < m$, то получим, что число узлов порядка $m$ оказывается не меньше числа заданных узлов, т. е. все концевые узлы, определяемые в соответствии с неравенством данной теоремы, могут быть получены и использованы при построении оптимального кода.

**Лемма.** *Средняя длина кода с основанием $K$ равна оптимальной длине в случае, если выполняется условие $p(x_{0i}) = K^{-a} = (1/K)^a$, где $1 \leqslant i \leqslant M$, $a$ — целое число.*

Учитывая условие $p(x_{0i}) = K^{-a}$ при разложении вероятностей появления сообщений в $K$-ичные числа, можно записать, что $p(x_{0i}) = K^{-n_i}$. Тогда средняя длина получаемого кода

$$n_{\text{ср}} = \sum_{i=1}^{M} n_i p(x_{0i}) = \sum_{i=1}^{M} n_i K^{-n_i}.$$

Оптимальная длина кода

$$n_{\text{опт}} = H(X_0)/H_{\max}(X) = -\sum_{i=1}^{M} p(x_{0i}) \log_2 p(x_{0i})/\log_2 K =$$
$$= -\sum_{i=1}^{M} K^{-n_i} \log_2 K^{-n_i}/\log_2 K = \sum_{i=1}^{M} n_i K^{-n_i}.$$

Таким образом, получаем, что $n_{\text{ср}} = n_{\text{опт}}$, т. е. построенный код является оптимальным. Рассмотрим методы построения оптимальных кодов.

*Метод дополнительных групп* заключается в том, что строится кодовое дерево, причем начало построения идет от концевых узлов, а кодовые комбинации записываются как обозначения концевых узлов кодового дерева. Дерево строится в следующей последовательности:

1) все сообщения располагаются в порядке убывания вероятностей их появления;

2) сообщения разбиваются на группы, причем в каждую группу объединяется $m$ сообщений, $2 \leqslant m \leqslant K$, $K$ — основание кода. Первая группа состоит из наименее вероятных сообщений;

3) каждая группа рассматривается как новое сообщение с вероятностью появления, равной сумме вероятностей появления сообщений, входящих в группу;

4) полученное новое сообщение располагается в соответствии с убыванием вероятностей;

5) полученному кодовому дереву ставятся в соответствие индексы от 0 до $K-1$.

Рассмотрим примеры построения оптимальных кодов.

а) Построение двоичного оптимального кода. Пусть имеется множество, состоящее из шести сообщений с вероятностями появления, равными 0,25; 0,25; 0,2; 0,15; 0,1; 0,05. Пользуясь изложенной выше методикой, построение оптимального кода покажем на рис. 36.

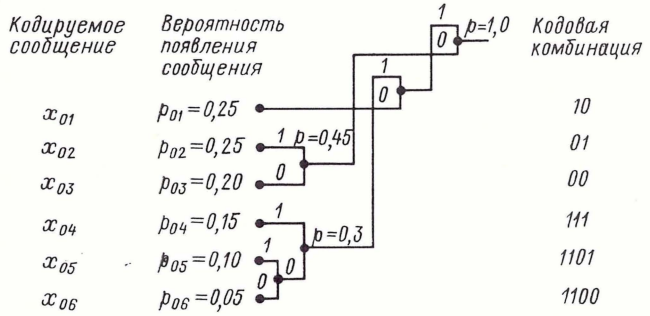


Рис. 36.

Проверим код на оптимальность. Средняя длина кода $n_{ср} = \sum_{j=1}^{M} n_j p(x_{0j})$. Оптимальная длина кода $n_{опт} = H(X_0)/\log_2 K = H(X_0)$. Для данного кода $n_{ср} = 2,45$, $n_{опт} = 2,42$. Если $n_{ср} - n_{опт} = \varepsilon \to 0$, то код оптимальный. Если $n_{ср} > n_{опт}$, то код избыточный. В данном случае избыточность вредна, так как не используется для обнаружения или исправления ошибок.

б) Построение троичного оптимального кода. Пусть имеется множество, состоящее из шести сообщений, с вероятностями появления 0,25; 0,25; 0,20; 0,15; 0,10; 0,05. Построение оптимального кода показано на рис. 37. Длина оптималь-

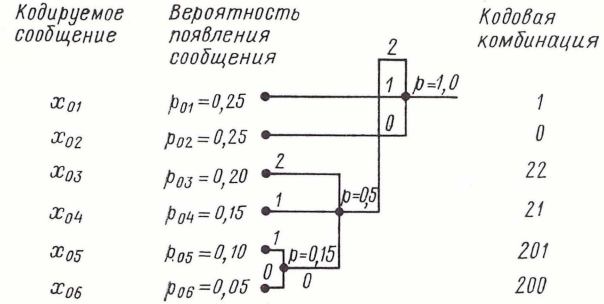


Рис. 37.

ного троичного кода $n_{опт} = H(X_0)/\log_2 3$. Подставляя значения в формулу, для данного случая получаем $n_{ср} = 1{,}65$, $n_{опт} = 1{,}58$. Число элементов в троичном коде оказывается меньше, чем в соответствующем двоичном, т. е. каждый элемент троичного кода является более информативным,

*Графический метод* является отображением кодового дерева, построение которого начинается с вершины. Последовательность построения следующая:

1) все события располагаются в виде точек на горизонтальной оси в порядке убывания вероятностей их появления;

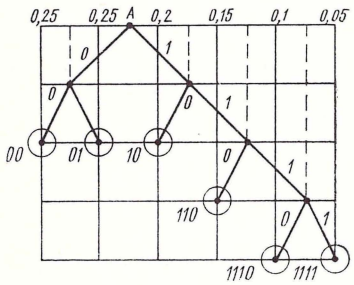


Рис. 38.

2) события разбиваются на $K$ групп, где $K$ — основание кода, и с данной точки начинается построение кодового дерева;

3) сообщения располагаются в концевых узлах кодового дерева, ветвям присваиваются индексы от 0 до $K-1$.

Пример построения двоичного оптимального кода. Пусть имеется множество, состоящее из шести сообщений, с вероятностями появления 0,25; 0,25; 0,20; 0,15; 0,10; 0,05. Пользуясь изложенной выше методикой, построим оптимальный код (рис. 38). При выборе начала кодового дерева ось вероятностей разбивается так, чтобы суммы вероятностей слева и справа при $K=2$ были равны.

Код, построенный графическим методом, обладает теми же свойствами, что и код, построенный методом дополнительных групп.

## § 3. Кодирование дискретной информации в каналах связи с шумом

При кодировании дискретной информации в канале связи с шумом основной задачей является построение кода, который обеспечивает скорость передачи информации, приближающуюся к пропускной способности канала связи, если вероятность ошибки не превышает заданного значения, или построение кода, обеспечивающего минимальную вероятность ошибки, если скорость передачи информации не меньше некоторого заданного значения.

К. Шенноном для дискретного канала связи с шумом доказано, что при производительности дискретного источника, не превышающей пропускную способность канала связи, можно так закодировать сообщения, что они будут переданы по каналу со сколь угодно малой вероятностью ошибки.

Неформальное доказательство состоит в следующем [28]: если множество сообщений кодируется последовательностями длиной $n$, то можно допустить, что передаваемые сообщения состоят из $2^{nH(X)}$ последовательностей, которые преобразуются на приемной стороне в $2^{nH(Y)}$ последовательностей, где $H(X)$, $H(Y)$ — энтропии источника на входе и на выходе канала связи.

Каждая последовательность на выходе канала связи с шумом является следствием примерно $2^{nH(X/Y)}$ последовательностей, поступивших на вход. Соответственно каждая последовательность на входе порождает $2^{nH(Y/X)}$ последовательностей на выходе канала связи. Пропускная способность канала связи с шумом

$$C = H(Y) - H(Y/X) = H(X) - H(X/Y).$$

Рассмотрим источник сообщений с производительностью $\nu \leqslant C$. Источник формирует $2^{n\nu}$ последовательностей длиной $n$. Оценим вероятность того, что никакая последовательность символов не была использована для кодирования сообщений. Эта вероятность равна

$$[1 - 2^{-nH(X)}]^{2^{n\nu}} \approx 1 - 2^{n[\nu - H(X)]}.$$

Вероятность того, что при приеме ни одна из $2^{nH(X/Y)}$ порождающих последовательностей не была выбрана неверно, составит

$$\{1 - 2^{n[\nu - H(X)]}\}^{2^{nH(X/Y)}} \approx 1 - 2^{-n(C-\nu)}$$

при условии, что для канала связи с шумом $C = H(X) - H(X/Y)$. Отсюда вероятность ошибки при передаче информации по каналу связи с шумом $P_0 = 2^{-n(C-\nu)}$. При $\nu \leqslant C$ и $n \to \infty$ $P_0 \to 0$. Следует отметить, что приведенное соотношение является асимптотическим и не может быть использовано для расчета вероятности ошибки при малых $n$.

Обсудим полученный результат. Пусть имеется двоичный симметричный канал с шумом. Пропускная способность этого канала $C = 1 + p \log_2 p + (1-p) \log_2 (1-p)$; $C_0 = vC$, где $v$ — производительность источника.

Будем считать, что вероятность трансформации есть функция $p(v/C_0)$, тогда [19]

$$1/(v/C_0) = 1 + p(v/C_0) \log_2 p(v/C_0) + \\ + [1 - p(v/C_0)] \log_2 [1 - p(v/C_0)].$$

На рис. 39 изображена зависимость $p(v/C_0)$. Из теоремы Шеннона следует, что если $v > C_0$, то всегда существует конечная вероятность ошибки и никаким кодом ее нельзя исключить. Поэтому возникает задача реализовать заданную вероятность ошибки $[P_0]_{\text{доп}}$. Необходимо обеспечить вероятность ошибки $P_0$ не больше $[P_0]_{\text{доп}}$, и для этого следует перейти к избыточному кодированию.

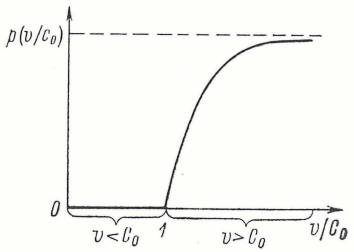

Рис. 39.

Пусть передается $M$ сообщений и используется код с основанием $K$, тогда $M = K^m$, где $m$ — число информационных элементов в коде. Избыточность — это некоторое дополнительное число элементов в коде, которое используется с целью обнаружения или исправления ошибок, вызванных действием помехи. Избыточность можно оценить коэффициентом И:

$$И = 1 - H(X_0)_m / H(X_0)_n,$$

где $H(X_0)_m$ — энтропия передаваемых сообщений, $H(X_0)_n$ — энтропия всего множества сообщений, формируемых кодом длиной $n$.

Пусть в коде $n$ элементов, причем $n = m + k$, где $k$ — число контрольных элементов. Из множества $M' = K^n$ выделим множество $M = K^m$, тогда $M' \to H(X_0)_n$, $M \to H(X_0)_m$. Если сообщения равновероятны, то $H(X_0)_n = \log_2 M'$, $H(X_0)_m = \log_2 M$ и

$$И = 1 - m \log_2 K / n \log_2 K = 1 - m/n = k/n.$$

Следовательно, при равновероятных сообщениях избыточность определяется долей контрольных символов в коде.

Рассмотрим формальную методику введения избыточности в код с целью обнаружения или обнаружения и исправления ошибок. В качестве математического аппарата, позволяющего описывать систематические коды, может быть использована теория матриц. Тогда для неизбыточного кода отображение можно получить с помощью квадратной единичной матрицы размерностью $m$:

$$I = \left\| \begin{matrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{matrix} \right\|.$$

Путем сложения по mod 2 в различных сочетаниях строк между собой получаем кодовые комбинации неизбыточного

двоичного кода длиной $m$. Например, если матрица имеет размерность $m=2$: $I=\begin{Vmatrix}1 & 0\\ 0 & 1\end{Vmatrix}$, то число возможных комбинаций $M=2^2-1=3$. Комбинации имеют вид 10, 01, 11. Таким образом, на основании единичной матрицы нетрудно построить всевозможные комбинации двоичного неизбыточного кода.

Для введения избыточности необходимо ввести $k$ контрольных символов. Построим некоторую производящую матрицу, которая позволит путем сложения отдельных строк в различных сочетаниях получить всевозможные комбинации избыточного кода [26]:

$$A=\begin{Vmatrix}1 & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1k}\\ 0 & 1 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2k}\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & 1 & a_{m1} & a_{m2} & \dots & a_{mk}\end{Vmatrix}.$$

Левая часть матрицы представляет собой единичную квадратную матрицу размерностью $m$, правая часть — дополнительную матрицу, содержащую $m$ строк и $k$ столбцов. Значения элементов, входящих в строки дополнительной матрицы, можно найти путем определенных формальных действий над элементами единичной матрицы.

Если избыточный код строится с кодовым расстоянием $d$, то строки дополнительной матрицы должны содержать $d-1$ единиц (одну единицу содержит строка единичной матрицы, стоящей слева). Различие между строками дополнительной матрицы должно быть не менее, чем в $d-2$ разрядах (различие в двух разрядах имеют строки единичной матрицы).

Для создания алгоритмов кодирования и декодирования используется проверочная матрица. Она строится следующим образом. К единичной матрице размерностью $k$ слева приписывают матрицу, которая содержит $m$ столбцов и $k$ строк, причем строка приписываемой матрицы есть столбец дополнительной матрицы в матрице $A$:

$$H=\begin{Vmatrix}a_{11} & a_{21} & \dots & a_{m1} & 1 & 0 & \dots & 0\\ a_{12} & a_{22} & \dots & a_{m2} & 0 & 1 & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ a_{1k} & a_{2k} & \dots & a_{mk} & 0 & 0 & \dots & 1\end{Vmatrix}.$$

Если в любой строке этой матрицы имеется некоторый набор единичных элементов, то они будут соответствовать тем символам кода, сумма которых по mod 2 в любой кодовой комбинации должна быть равна нулю.

Таким образом, используя элементы проверочной матрицы, можно получить уравнения проверки, число которых соответствует числу строк проверочной матрицы. На основе уравнений проверки можно определить значения символов в контрольных разрядах и, имея лишь неизбыточный код длиной $m$, с учетом полученных контрольных соотношений найти значения символов на контрольных позициях, т. е. перейти к избыточному коду. Эти же операции по проверке выполняются и на приемной стороне. Тогда в итоге проверки возникает результат в виде кодовой комбинации, которая содержит $k$ разрядов. Полученное $k$-разрядное двоичное число является опознавателем (синдромом), в соответствии с которым находят место и характер ошибки.

Рассмотрим пример построения некоторого кода, отображающего десятичный разряд, т. е. способного отобразить десять возможных знаков, и потребуем, чтобы код исправлял одну ошибку. Кодовое расстояние $d = 3$. Поскольку необходимо иметь $M \geqslant 10$, то число информационных позиций в коде $m = 4$. Нетрудно показать, что число проверочных символов равно трем. Тогда производящая матрица будет иметь следующий вид:

$$A = \left\| \begin{matrix} 1000 & 011 \\ 0100 & 101 \\ 0010 & 110 \\ 0001 & 111 \end{matrix} \right\|.$$

На основании данной матрицы можно построить комбинации кода, число которых равно 16. Расстояние между комбинациями, получаемыми путем суммирования отдельных строк матрицы, равно 3. Строим проверочную матрицу

$$H = \left\| \begin{matrix} 0111100 \\ 1011010 \\ 1101001 \end{matrix} \right\|,$$

из которой нетрудно получить уравнения по определению значений символов на контрольных позициях. Используя единицы первой строки матрицы $H$, находим уравнение $x_5 = x_2 \oplus x_3 \oplus x_4$. Используя элементы второй строки, получаем $x_6 = x_1 \oplus x_3 \oplus x_4$. Используя элементы третьей строки, имеем $x_7 = x_1 \oplus x_2 \oplus x_4$. На приемной стороне необходимо найти синдром с учетом всех возможных элементов кода, как информационных, так и контрольных. Тогда первая проверка включает в себя элементы $\alpha_1 = x_2 \oplus x_3 \oplus x_4 \oplus x_5$, вторая $\alpha_2 = x_1 \oplus x_3 \oplus x_4 \oplus x_6$, третья $\alpha_3 = x_1 \oplus x_2 \oplus x_4 \oplus x_7$. Выполняя эти проверки, находим трехразрядное двоичное число, кото-

рое является опознавателем; сопоставляя значение опознавателя с характером возникшей ошибки, можно всегда найти соответствие между значением контрольного числа и местом появляющейся ошибки.

Для рассматриваемого кода кодовые комбинации имеют вид: 0001111, 0010110, 0011001, 0100101, 0101010, 0110011, 0111100, 1000011, 1001110, 1010101. Выбирая произвольную комбинацию и вводя последовательно одиночные ошибки на первой — седьмой позициях, найдем опознаватели путем определения символов $\alpha_1$, $\alpha_2$, $\alpha_3$ по вышеприведенным уравнениям. Опознаватели имеют следующий вид.

| Место ошибки | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Опознаватель | 011 | 101 | 110 | 111 | 100 | 010 | 001 |

При наличии ошибки в переданной комбинации выполняем на приемной стороне $k$ проверок и по полученному опознавателю находим место ошибки.

Возможно такое построение избыточного кода, когда соответствие имеет простейший характер. Тогда опознаватель представляет собой некоторое двоичное число, позволяющее при переводе в десятичную форму указать номер позиции, в которой произошла ошибка. Коды, подчиняющиеся этому условию, получили название кодов Хэмминга. Особенностью их является то, что контрольные позиции располагаются в коде по закону $2^i$. При этом в первое проверочное уравнение включаются позиции, номера которых в двоичном представлении содержат в младшем разряде единицу, т. е. $\alpha_1 = x_1 \oplus x_3 \oplus x_5 \oplus \ldots$ Соответственно во второе уравнение проверки включаются те позиции, номера которых в двоичном представлении содержат единицу во втором разряде, тогда $\alpha_2 = x_2 \oplus x_3 \oplus x_6 \oplus \ldots$, и т. д. Проверочную матрицу получаем с учетом того, что элементы единичной матрицы занимают первый, второй, четвертый столбцы, а остальные столбцы получаются непосредственно из строк дополнительной матрицы, входящей в образующую матрицу $A$:

$$H = \left\| \begin{matrix} 101 & 0101 \\ 011 & 0011 \\ 000 & 1111 \end{matrix} \right\|,$$

т. е. имеется возможность передать, так же как и ранее, $M = 16$ сообщений при $m = 4$, $k = 3$, $n = 7$. Код Хэмминга реализует предел избыточности для независимых ошибок в симметричном канале связи, что будет рассмотрено далее.

## § 4. Информационные пределы избыточности

При введении избыточности улучшаются обнаруживающие и исправляющие свойства кода. Количественно избыточность характеризуется числом кодовых переходов, однако увеличение избыточности приводит к нерациональному использованию кода, что выражается в снижении скорости передачи информации. Существует минимально необходимая избыточность, которая связана с заданным числом переходов $d$. Нижний предел избыточности называется *информационным* пределом, найдем его.

*Информационный предел избыточности для канала с независимыми ошибками.* Пусть имеется код с основанием $K=2$, который исправляет ошибки, причем $r=s$ и $d=2s+1$. В таком коде возможны два исхода: прохождение сообщения с вероятностью $P_{\text{пр}}$ (область $A$) и трансформация сообщения с вероятностью $P_{\text{тр}}$ (область $B$). На рис. 40 показаны эти области приема кодовой комбинации $x_{0j}$.

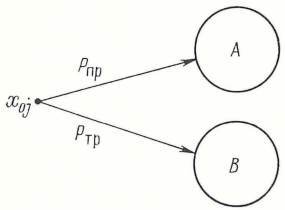

Рис. 40.

Пусть послана кодовая комбинация $x_{0j}(x_{1j}\ x_{2j}\ \ldots\ x_{nj})$, на которую воздействует помеха $z_{0j}\,(z_{1j}\,z_{2j}\ \ldots\ z_{nj})$, где $x_{1j}$ — первый символ кодовой комбинации, $x_{nj}$ — $n$-й символ. Наличие ошибки отображается появлением на $i$-й позиции $z_{ij}$. На приемной стороне возникает комбинация кода $y_{0i}=x_{0j}+z_{0j}$, которая может существенно отличаться от посланной кодовой комбинации $x_{0j}$, поэтому необходимо иметь определенный объем контрольной информации [13] $I_{\text{к}}=I_{\text{к}A}+I_{\text{к}B}$, $I_{\text{к}A}$ — количество информации, необходимое для оценки попадания сообщений в область $A$, $I_{\text{к}B}$ — количество информации, связанное с попаданием сообщений в область $B$:

$$I_{\text{к}}=-\sum_{j\in A} p_j \log_2 p_j-\sum_{j\in B} p_j \log_2 p_j,$$

при этом $p_j$ — вероятность попадания в точку $j$ для принимаемого кодового множества; $I_{\text{к}}$ — это информация, которая должна быть доставлена с передающей стороны на приемную. Поскольку информация передается в элементах кода, то для $K$-ичного кода в одном символе может передаваться максимальное количество информации $H_{\max}(X)=\log_2 K$.

Будем считать, что контрольная информация передается набором контрольных символов, тогда минимально возможное число контрольных символов $k_{\min}=I_{\text{к}}/\log_2 K$ называется *информационным пределом избыточности*. Данный минимум может быть достигнут, когда в коде все расстояния между ком-

бинациями одинаковы, т. е. код является эквидистантным. Необходимо, чтобы символы кода появлялись с одинаковой вероятностью с целью обеспечения максимального количества информации в одном символе. Для определения $k_{\min}$ необходимо знать модель помех в канале связи.

Рассмотрим двоичный канал с независимыми ошибками, когда ошибки не группируются в пакеты и появление предыдущей ошибки не влечет за собой последующую. Тогда действует биноминальный закон распределения ошибок, и вероятность возникновения ошибки кратности $i$ данного вида

$$p(i) = p^i (1-p)^{n-i},$$

где $p$ — вероятность возникновения ошибки. Определим количество контрольной информации:

$$I_к = -\sum_{j \in A} M(i)\, p(i) \log_2 p(i) - P_{тр} \log_2 P_{тр},$$

где $M(i)$ — число вариантов ошибок кратности $i$, тогда

$$I_к = -\sum_{i=0}^{s} C_n^i p^i (1-p)^{n-i} \log_2 \left[p^i (1-p)^{n-i}\right] - P_{тр} \log_2 P_{тр}.$$

Вторым членом можно пренебречь, так как не исправляем ошибки трансформации:

$$I_к = -\sum_{i=0}^{s} C_n^i p^i (1-p)^{n-i} \log_2 \left[p^i (1-p)^{n-i}\right].$$

Рассмотрим частный случай, когда возникновение любой корректируемой ошибки и отсутствие ошибок имеют одинаковую вероятность: $p^i (1-p)^{n-i} = p_1$, $I_к = -\sum_{i=0}^{s} C_n^i p_1 \log_2 p_1$.

Поскольку общее число возможных видов событий кратностей ошибок, не приводящих к трансформации, $s+1$, то $\sum_{i=0}^{s} C_n^i p_1 = 1$ и $p_1 = 1/\sum_{i=0}^{s} C_n^i$, следовательно:

$$I_к = -\sum_{i=0}^{s} C_n^i p_1 \log_2 p_1 = \log_2 \sum_{i=0}^{s} C_n^i.$$

В этом случае $k_{\min} = \log_2 \sum_{i=0}^{s} C_n^i$ называется *нижним пределом Хэмминга*. Необходимое число контрольных элементов $k \geqslant k_{\min} = \log_2 \sum_{i=0}^{s} C_n^i$, отсюда $2^k \geqslant \sum_{i=0}^{s} C_n^i$. Пусть код содержит $n$ элементов, $n = m + k$, $M = 2^m$, следовательно, $2^{n-m} \geqslant \sum_{i=0}^{s} C_n^i$, или $2^m \leqslant 2^n / \sum_{i=0}^{s} C_n^i$. Это выражение является практической формулой построения кода при $d = 2s+1$. Если $s = 1$, $d = 3$, $M \leqslant 2^n/n + 1$, получаем код, исправляющий одну ошибку. Легко показать, что избыточность кодов Хэмминга, т. е. число контрольных позиций, уменьшается с увеличением $M$, причем при $M \to \infty$ $k/n \to 0$.

*Информационный предел избыточности для канала с пакетными ошибками.* Пусть имеется канал связи, в котором

возникают пакетные ошибки. Под пакетной ошибкой кратности $i$ будем понимать ошибку в коде, имеющем длину $i$. Такие ошибки обычно возникают в коммутируемых телефонных линиях связи при кратковременном разрыве коммутируемых цепей. Оценим информационный предел избыточности. Для исправления $s$ ошибок введем некоторый объем контрольной информации:

$$I_к = -\sum_{i=0}^{s} M(i)\, p(i) \log_2 p(i) - P_{тр} \log_2 P_{тр},$$

где $M(i)$ — число разновидностей ошибок с длиной пакета $i$. Можно показать [13], что

$$M(i) = \begin{cases} 1 & \text{при } i = 0, \\ n & \text{при } i = 1, \\ (n - i + 2)\, 2^{i-2} & \text{при } i \geqslant 2. \end{cases}$$

Полагая, что ошибки трансформации не исправляются, получаем:

$$I_к = -p(0) \log_2 p(0) - np(1) \log_2 p(1) - \\ - \sum_{i=2}^{s} (n - i + 2)\, 2^{i-2} p(i) \log_2 p(i).$$

Если модель помех такова, что $p(i)$ не зависит от $i$, т. е. $p(0) = p(1) = p(i) = p_1$, то

$$I_к = -p_1 \log_2 p_1 - np_1 \log_2 p_1 - \sum_{i=2}^{s} (n - i + 2)\, 2^{i-2} p_1 \log_2 p_1.$$

Найдем $p_1$ из условия $\sum_{i=0}^{s} M(i)\, p(i) = 1$:

$$p_1 = 1/1 + n + \sum_{i=2}^{s} (n - i + 2)\, 2^{i-2}.$$

Количество контрольной информации

$$I_к = \log_2 \left[1 + n + \sum_{i=2}^{s} (n - i + 2)\, 2^{i-2}\right].$$

Количество контрольных элементов в двоичном коде

$$k \geqslant k_{\min} = I_к / \log_2 K = I_к.$$

Реализация данного информационного предела возможна при равновероятных пакетных ошибках различной длины.

## § 5. Информационные процессы при кодировании и декодировании

Кодирование можно представить как переход от множества сообщений $X_0$ к некоторому множеству символов $V$, при котором сохраняется однозначное соответствие. Различают непрерывное и блочное кодирование, в первом случае поток

сообщений кодируется общим кодом и не разделяется на отдельные части, во втором — каждое сообщение отображается кодом с постоянной длиной.

Рассмотрим процесс блочного кодирования и декодирования (рис. 41). В канал связи поступают последовательно символы $x_{ji}$ из множества $V = X^n$. Символы располагаются в блоках длиной $n$. Возникает задача определения начала и конца каждого блока. Если $n$ постоянно, то достаточно определить лишь начало блока. При схемной реализации процесс установления начала получил название синхронизации. Будем считать, что синхронизация обеспечена, и остановимся на процессе декодирования. Декодированием называется переход от множества символов $U = Y^n$ к множеству принимаемых сообщений $Y_0$. В процессе декодирования обеспечивается реализация внесенной избыточности и основными исходами являются правильный прием, трансформация и защитный отказ. Соотношения между вероятностями событий $P_{\text{пр}}$, $P_{\text{тр}}$, $P_{\text{зо}}$ определяются не только каналом связи, но и методом приема. Существуют следующие наиболее характерные способы декодирования:

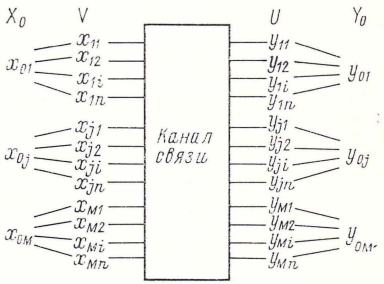

Рис. 41.

1. Известна статистика всех передаваемых сообщений, т. е. известен закон распределения вероятностей появления сообщений: $P(x_{01}, x_{02}, \ldots, x_{0M}) = P(X_0)$, известна статистика появления символов в канале связи: $P(x_1, x_2, \ldots, x_K) = P(X)$. На приемной стороне оценивается вероятность передачи заданного сообщения $x_{0j}$ при появлении некоторой комбинации $y_{0i}$ и используется формула Байеса, т. е. оценивается вероятность того, что передано сообщение $x_{0j}$, если принято сообщение $y_{0i}$:

$$p(x_{0j}/y_{0i}) = p(x_{0j})\, p(y_{0i}/x_{0j})/P(X_0).$$

Просчитываем $M$ вариантов и определяем

$$p(x_{01}/y_{Ci}),\ p(x_{02}/y_{Ci}),\ \ldots,\ p(x_{0M}/y_{0i}),$$

затем выбираем наибольшую из этих вероятностей и получаем, что $P_{\max}(x_{0j}/y_{0i})$ соответствует $x_{0j}$.

Вероятность ошибки в данном методе $P_0 = 1 - P_{\max}(x_{0j}/y_{0i}) = P_{0\min}$, откуда видно, что для сообщения с индексом $j$ решение получено с минимальной вероятностью ошибки. Поэтому этот критерий получил название *критерия минимума вероятности ошибки.*

2. Статистика передаваемых сообщений неизвестна. В этом случае весь анализ проводится только на основании информации, получаемой с выхода канала связи. Прием осуществляется по *критерию максимума правдоподобия*. Оцениваем на приемной стороне статистику появления символов по каждому сигнальному признаку: $P(y_1, y_2, \ldots, y_K) = P(Y)$, где $y_1, y_2, \ldots, y_K$ — символы различных сигнальных признаков, например 0, 1, 2, ...

Считаем, что $P(Y)$ хранится в памяти, и анализируем принимаемые сообщения. Пусть принята некоторая последовательность $(y_{1i}, y_{2i}, \ldots, y_{ni})$, которая отображает сообщение $y_{0i}$, нам неизвестное. Определяем $p(y_i/x_j)$ и находим величину

$$d_{ij} = \ln [p(y_i/x_j)/P(Y)],$$

которая показывает отклонение принятого символа от переданного; $d = \sum_{i,j=1}^{n} d_{ij}$ — расстояние между $x_{0j}$ и $y_{0i}$. Выбираем $d_{\min}$ для некоторого $j$, тогда оно соответствует максимальному правдоподобию конкретной кодовой комбинации.

Предположим, что используется один из методов приема информации. Введем кодирующее и декодирующее устройства в канал связи и предположим, что в нем действует помеха, не позволяющая однозначно выявить сообщения. Тогда количество информации, передаваемой через канал связи, можно установить из модели, показанной на рис. 42:

Рис. 42.

$$I(X_0, Y_0) = H(Y_0) - H(Y_0/X_0),$$

где $H(Y_0/X_0)$ — количество информации, теряемой под действием помехи, соответствует избыточной информации, которую нужно ввести. Если необходимо обнаруживать и исправлять ошибки, то эта величина есть количество информации, идущей на обнаружение и исправление ошибки:

$$H(Y_0/X_0) = H_0 + H_и,$$

где $H_0$ — энтропия двух событий. Одно из них отображает наличие ошибки с вероятностью $P_0$, а второе — отсутствие ошибки с вероятностью $1 - P_0$:

$$H_0 = -P_0 \log P_0 - (1 - P_0) \log_2 (1 - P_0).$$

Эта формула справедлива, если для каждого сообщения вероятность ошибки одинакова. Значение $H_0$ тем меньше, чем меньше вероятность ошибки.

Если сообщения имеют разные вероятности ошибки, то можно ввести понятие вероятности ошибки при приеме $i$-го сообщения:

$$P(e/y_{0i}) = \sum_{i=1}^{M} p(x_{0j})\, p(y_{0i}/x_{0j}) \text{ при } i \neq j.$$

Тогда средняя вероятность ошибки

$$P_0 = \sum_{i=1}^{M} p(y_{0i})\, P(e/y_{0i}).$$

Количество информации, связанное с обнаружением ошибки в $i$-м сообщении:

$$H(e/y_{0i}) = -P(e/y_{0i}) \log_2 P(e/y_{0i}) - \\ -[1 - P(e/y_{0i})] \log_2 [1 - P(e/y_{0i})],$$

тогда среднее количество информации, необходимое для обнаружения ошибки:

$$H_0 = \sum_{i=1}^{M} p(y_{0i})\, H(e/y_{0i}).$$

Найдем количество информации, необходимое для исправления ошибки. Для этого рассмотрим передачу некоторого сообщения. На приемной стороне попадаем в множество, содержащее $M-1$ сообщений. Если переходы между отдельными сообщениями равновероятны, то неопределенность состояния $H = \log_2 (M-1)$, но так как она имеет место, когда есть ошибка с вероятностью $P_0$, то $H_и = P_0 \log_2 (M-1)$, а

$$H(Y_0/X_0) = -P_0 \log_2 P_0 - (1 - P_0) \log_2 (1 - P_0) + \\ + P_0 \log_2 (M-1).$$

Таким образом, рассмотрение пределов избыточности показывает:

1) для компенсации потерь информации в канале связи необходимо ввести избыточную информацию, определяемую из условия обнаружения и исправления ошибок;

2) полученное значение избыточности информации является минимальным и реализуется при оптимальном согласовании источника информации с каналом связи;

3) основное влияние на вероятность ошибки оказывает соотношение скорости и пропускной способности канала связи.

Рассмотрим канал связи, в котором передача осуществляется с использованием $K$-ичного кода, т. е. следует оценить необходимый уровень избыточной информации и возможности повышения скорости передачи в таких каналах. Сообщение $x_{0i}$ может попасть либо в область правильного приема $A$, либо под действием помех в область неправильного приема $B$.

Код построен так, что все обнаруживаемые ошибки исправляются, число переходов $d = 2s + 1$, $r = s$. Найдем минимально необходимое количество контрольной информации из условия $I_k = H_A + H_B$. Так как ошибки трансформации не исправляем, т. е. $H_B = 0$, то

$$I_к = H_A = -\sum_{i \in A} p(i) \log_2 p(i).$$

Введем понятие ошибки $i$-го типа. По месту положения ошибки подчиняются биноминальному закону распределения. Ошибка $i$-го типа — это ошибка, связанная с подавлением или воспроизведением $i$ элементов в коде. Ошибка $i$-й кратности данного вида может быть воспроизведена с вероятностью $p(i) = p^i (1-p)^{n-i}$. Число возможных вариантов ошибок $i$-го типа $M(i) = C_n^i (K-1)^i$. Например, пусть имеется троичный код 1012. Найдем число возможных ошибок кратности два. По месту положения число ошибок составит $C_n^2$. По характеру на каждом из местоположений возможно $2^2$ сочетаний. Всего получаем 24 возможных варианта. Количество контрольной информации

$$I_к = H_A = -\sum_{i=0}^{s} C_n^i (K-1)^i p^i (1-p)^{n-i} \log_2 \left[p^i (1-p)^{n-i}\right].$$

Если модель помехи такова, что появление любой разновидности ошибки $i$-го типа происходит с одинаковой вероятностью $p_1 = p^i (1-p)^{n-i}$, тогда

$$I_к = -\sum_{i=0}^{s} C_n^i (K-1)^i p_1 \log_2 p_1.$$

С учетом возможных ошибок трансформации $\sum_{i=0}^{n} C_n^i (K-1) p_1 = 1$, откуда $p_1 = 1/\sum_{i=0}^{n} C_n^i (K-1)^i$. Следовательно,

$$I_к = -\left[\sum_{i=0}^{s} C_n^i (K-1)^i / \sum_{i=0}^{n} C_n^i (K-1)^i\right] \log_2 1/\sum_{i=0}^{n} C_n^i (K-1)^i.$$

Минимальное число контрольных позиций

$$k_{\min} = I_к / H_{\max}(X) = I_к / \log_2 K,$$

где $H_{\max}(X)$ — максимальная энтропия источника; $I_к$ и $k_{\min}$ дают информационный предел избыточности для $K$-ичного канала с независимыми ошибками.

Рассмотрим частный случай: появление ошибок с кратностью больше $s$ маловероятно, т. е. выбрана соответствующая способность кода, тогда

$$I_к = \log_2 \sum_{i=0}^{s} C_n^i (K-1)^i,$$

$$k_{\min} = \left[\log_2 \sum_{i=0}^{s} C_n^i (K-1)^i\right] / \log_2 K.$$

Необходимое число контрольных символов $k \geqslant k_{\min}$:

$$k \geqslant \log_K \sum_{i=0}^{s} C_n^i (K-1)^i.$$

Тогда $K^k \geqslant \sum_{i=0}^{s} C_n^i (K-1)^i$. Если в коде имеется $k$ контрольных и $m$ информационных символов, то $n = k + m$ и $M = K^m = K^n / \sum_{i=0}^{s} C_n^i (K-1)^i$. Код позволяет исправить $s$ ошибок, т. е. $d = 2s + 1$. При $s = 1$ $M = K^m \leqslant K^n / 1 + n(K - 1)$, $d = 3$.

Вероятность ошибки в $K$-ичном канале можно оценить, если известен критерий выявления сообщений. Пусть используется критерий максимума правдоподобия, т. е. на приемной стороне определяется $d_{ij} = \ln [p(y_i/x_j)/P(Y)]$. Далее устанавливается кодовое расстояние $d = \sum_{i,j=1}^{n} d_{ij}$. Величина $d_{ij}$ может быть определена двояко.

1. Сигнальные признаки не могут переходить друг в друга под действием одиночной ошибки, тогда считают, что $d_{ij} = 2$, и расстояние между кодовыми комбинациями 1012, 2011 составит $d = 4$.

2. Между сигнальными признаками один переход. Тогда расстояние между кодовыми комбинациями 1012, 2011 $d = 2$. Это различие определяется типом сигнального признака. В первом случае в качестве сигнального признака используются частота, фаза, полярность, во втором — амплитуда.

Вероятность ошибки можно определить следующим образом: пусть расстояние между соседними колебаниями $i$ и $j$ есть $d_0$. Тогда

$$P_{oi} = p\{d(u_i, v) > d_o\},$$

где $u_i$ — принятая кодовая комбинация; $v$ — множество комбинаций, заданных на входе канала связи. Усредняя вероятность ошибки по всем возможным комбинациям, получаем

$$P_o = \sum_{i=1}^{M} p(u_i) P_{oi},$$

где $M$ — число передаваемых сообщений. Обычно необходимо получить $P_0 \leqslant [P_0]_{\text{доп}}$. При этом вероятность ошибки распадается на составляющие, т. е. $P_0 = P_{\text{тр}} + P_{\text{зо}}$ и $P_{\text{тр}} \leqslant [P_{\text{тр}}]_{\text{доп}}$. Перераспределение между $P_{\text{тр}}$ и $P_{\text{зо}}$ осуществляется на приемной стороне путем соответствующего конструирования декодирующего устройства. На входе канала обычно выбирают значение $d = r + s + 1$. Чтобы обеспечить условие $P_{\text{тр}} \leqslant [P_{\text{тр}}]_{\text{доп}}$, необходимо выбрать $r > s$, тогда увеличивается обнаруживающая способность кода и возрастает вероятность $P_{\text{зо}}$ при одновременном уменьшении $P_{\text{тр}}$.

## § 6. Граничное значение вероятности ошибки при производительности источника, превышающей пропускную способность канала связи

Пусть имеется некоторый канал связи с пропускной способностью $C$ и по нему передается $n$-элементный код. Положим, что производительность источника $H(X_0) > nC$, где $H(X_0)$ — количество информации, выдаваемое в некотором усредненном сообщении, $n$ — длина кода, $C$ — максимально возможное количество информации, которое может передать один элемент кода в заданном канале связи, т. е. пропускная способность канала связи.

**Теорема.** *Если производительность источника превышает пропускную способность канала связи, т. е.* $H(X_0) > nC$, *то невозможна передача информации с бесконечно малой вероятностью ошибки. Нижняя граница вероятности ошибки определяется из уравнения*

$$-P_0 \log_2 P_0 - (1 - P_0) \log_2 (1 - P_0) + P_0 \log_2 (M - 1) = \\ = H(X_0) - nC.$$

Доказательство. Количество информации, передаваемой через канал связи, определяется формулой

$$I(X_0, Y_0) = H(Y_0) - H(Y_0/X_0),$$

где $H(Y_0/X_0)$ — количество избыточной информации, которое должно быть внесено в передаваемый сигнал. Так как через канал связи нельзя передавать количество информации, превышающее пропускную способность, то

$$I(X_0, Y_0) \leqslant nC.^*$$

Будем считать, что вероятность ошибки одинакова для каждого сообщения, тогда

$$H(Y_0/X_0) = -P_0 \log_2 P_0 - (1 - P_0) \log_2 (1 - P_0) + \\ + P_0 \log_2 (M - 1).$$

Подставляя это выражение в (*), получаем

$$H(Y_0) + P_0 \log_2 P_0 + (1 - P_0) \log_2 (1 - P_0) - \\ - P_0 \log_2 (M - 1) \leqslant nC,$$

откуда $-P_0 \log_2 P_0 - (1 - P_0) \log_2 (1 - P_0) + P_0 \log_2 (M - 1) \geqslant \geqslant H(Y_0) - nC$.

Поскольку $H(X_0) \leqslant H(Y_0)$, то вместо $H(Y_0)$ можно подставить $H(X_0)$:

$$-P_0 \log_2 P_0 - (1-P_0)\log_2(1-P_0) + P_0 \log_2 (M-1) \geqslant H(X_0) - nC.$$

Проанализируем полученное выражение. Обозначим $P_0 = z$, где $z$ — промежуточная переменная, и рассмотрим функцию

$$f(z) = -z \log_2 z - (1-z)\log_2(1-z) + z \log_2 (M-1),$$

причем $f(z) > 0$ при $z > 0$.

Приравняем к нулю первую производную:

$$f'(z) = k\left[-\frac{z}{z} - \ln z + \frac{(1-z)}{(1-z)} + \ln(1-z) + \ln(M-1)\right] =$$
$$= k \ln\left[\frac{(M-1)(1-z)}{z}\right] = 0.$$

Решая это уравнение, находим $z = 1 - 1/M$. При полученном значении $z$ вероятность ошибки является минимальной: $f(z) > f(z = 1 - 1/M)$. Таким образом, вероятность ошибки не может быть ниже полученного минимального значения.

При объединении нескольких информационных потоков необходимо учитывать возможность превышения пропускной способности канала связи и невозможность сведения к минимуму вероятности ошибки. На практике согласование осуществляется с помощью коммутационных центров. Пусть имеются источники информации 1И, 2И, ..., $N$И и центральный пост (ЦП), куда должна передаваться информация (рис. 43). На пути следования информации установим коммутационный центр (промежуточный пункт, ПП) и объединим источники в реальную структуру. За промежуточным пунктом следует канал связи с пропускной способностью $C$, причем $nC > H(X_0)$, т. е. $nC > \sum_{i=1}^{N} C_i n_i$, и можно обеспечить бесконечно малую вероятность ошибки.

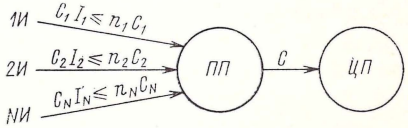


Рис. 43.

## § 7. Эквивалентная вероятность ошибки

Пусть имеется некоторое множество передаваемых сообщений и каждое сообщение в среднем несет количество информации, равное энтропии:

$$H(X_0) = -\sum_{j=1}^{M} p(x_{0j}) \log_2 p(x_{0j}).$$

Эквивалентной вероятностью ошибки $p_э$ является вероятность ошибки при передаче одной двоичной единицы информации [22]. Если сообщение передается с вероятностью ошибки $P_0$ и содержит $H(X_0)$ двоичных единиц информации, то вероятность прохождения сообщения

$$P_{пр} = 1 - P_0 = (1 - p_э)^{H(X_0)}.$$

Так как вероятность $p_э$ мала, то считаем $P_{пр} \approx 1 - p_э H(X_0)$, тогда $1 - P_0 \approx 1 - p_э H(X_0)$, $p_э \approx P_0/H(X_0)$.

Сравним различные коды по эквивалентной вероятности ошибки, если передача информации осуществляется по каналу с независимыми ошибками.

1) Двоичный неизбыточный код: $M = 2^n$, $d = 1$, $r = s = 0$. Так как неизбыточный код используется в канале без шума при равновероятности передаваемых сообщений, то $p(x_{0j}) = 1/M$. Энтропия сообщения при равновероятности событий $H(X_0) = \log_2 M = n$. Следовательно, $p_э = P_0/n = [1 - p(0)]/n$, где $p(0)$ — вероятность непопадания помех на код. Для пуассоновского закона распределения импульсной помехи

$$p(0) = \exp(-I_{сп}),\ p_э \approx [1 - \exp(-I_{сп})]/n \approx I_{сп}/n,$$

где $I_{сп}$ — интенсивность случайной импульсной помехи в канале связи [29].

Из формулы следует, что $p_э$ есть интенсивность помехи, приведенная к одному элементу кода.

2) $K$-ичный неизбыточный код: $M = K^n$, $d = 1$, $r = s = 0$. Энтропия сообщения $H(X_0) = \log_2 M = n \log_2 K$. Введем вероятность искажения одного элемента кода $p$. Тогда вероятность прохождения сообщения

$$P_{пр} = 1 - P_0 = (1 - p)^n = (1 - p_э)^{H(X_0)}.$$

Так как $(1 - p)^n = (1 - p_э)^{n \log_2 K}$, т. е. $(1 - p)^{1/\log_2 K} = 1 - p_э$, а $p$ мало, получаем

$$1 - p/\log_2 K \approx 1 - p_э, \text{ или } p_э \approx p/\log_2 K.$$

С увеличением основания кода при одной и той же эквивалентной вероятности ошибки можно допустить большую вероятность искажения элемента кода $p$ в канале связи, т. е. увеличение основания кода является полезным.

3) Двоичный избыточный код: $M \leqslant 2^n / \sum_{i=0}^{s} C_n^i$, $d = 2s + 1$. Эквивалентная вероятность ошибки $p_э \approx P_0/H(X_0)$. Так как в двоичном коде лишь $m$ позиций несут информацию, а остальные позиции контрольные, то

$$H(X_0) = \log_2 M = \log_2 2^m = m.$$

Вероятность ошибки в приеме сообщения

$$P_0 = 1 - \sum_{i=0}^{s} p(i) \approx p(s+1),$$

где $p(i)$ — вероятность попадания $i$ импульсов помехи на код. Поэтому

$$P_0 \approx I_{\text{сп}}^{(s+1)} \exp(-I_{\text{сп}})/(s+1)!,$$

$$\text{или } p_{\text{э}} \approx I_{\text{сп}}^{(s+1)} \exp(-I_{\text{сп}})/[(s+1)!\, m].$$

Значение $p_{\text{э}}$ уменьшается с увеличением длины информационной части кода, а также с увеличением числа исправляемых ошибок. Значение $s$ ограничивается сложностью схемной реализации.

4) $K$-ичный избыточный код: $M \leqslant K^n / \sum_{i=0}^{s} C_n^i (K-1)^i$; $d = 2s+1$. Эквивалентная вероятность ошибки $p_{\text{э}} \approx P_0 / H(X_0)$. Если сообщения равновероятны, то энтропия сообщения

$$H(X_0) = \log_2 M = \log_2 K^m = m \log_2 K.$$

Вероятность ошибки в приеме сообщения $P_0 = p(s+1)$,

$$p_{\text{э}} \approx p(s+1)/m \log_2 K = I_{\text{сп}}^{(s+1)} \exp(-I_{\text{сп}})/[(s+1)!\, m \log_2 K].$$

Для снижения эквивалентной вероятности ошибки следует увеличивать информационную часть кода, основание кода и число исправляемых ошибок. На практике достаточно просто реализуются троичные коды с исправлением одиночной ошибки и пятеричные с исправлением двух ошибок.

## *Глава V*

# НЕПРЕРЫВНАЯ ИНФОРМАЦИЯ В АСУ

Значительная часть передаваемой в АСУ информации является непрерывной и отображает физические процессы, представляемые в виде некоторого непрерывно меняющегося напряжения или тока.

Непрерывная информация может передаваться двумя способами:

1) непосредственной передачей непрерывного сигнала на незначительное расстояние, что означает выделение случайного сигнала на фоне помех и реализуется с помощью оптимальных фильтров;

2) дискретизацией непрерывной функции с последующим восстановлением ее на приемной стороне по принятым дискретным значениям.

Непрерывная информация может быть изображена графически, аналитически и спектрально. По графическому изображению при наличии ряда реализаций можно оценить статистику появления отдельных значений функции. Графическое задание полезно и особенно важно при отображении функции на приемной стороне.

Используя спектральное представление, непрерывную функцию можно записать в виде бесконечного ряда

$$f(t) = \sum_{k=-\infty}^{\infty} C_k \cos(k\omega_1 t - \varphi_k),$$

где $C_k$ — амплитуда, $\varphi_k$ — начальная фаза $k$-й составляющей, или в виде

$$f(t) = \sum_{k=-\infty}^{\infty} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t),$$

причем $$C_k = \sqrt{a_k^2 + b_k^2}, \quad \operatorname{tg} \varphi_k = b_k/a_k.$$

Весьма удобно представление функции в комплексной форме:

$$f(t) = \sum_{k=-\infty}^{\infty} C_k \exp(i2\pi kt/T),$$

где $C_k = (1/T)\int_{-T/2}^{T/2} f(t)\exp(-i\,2\pi kt/T)\,dt$, что справедливо для периодической функции с периодом $T$. Если функция непериодическая, то переходим к интегралу Фурье

$$f(t) = (1/2\pi)\int_{-\infty}^{\infty} s(\omega)\exp(i\omega t)\,d\omega,$$

и тогда плотность амплитудного спектра

$$s(\omega) = \int_0^{\infty} f(t)\exp(-i\omega t)\,dt.$$

## § 1. Дискретизация непрерывной информации

Одним из способов передачи непрерывной информации является ее дискретизация во времени, что позволяет перейти к дискретной информации и использовать рассмотренные выше способы кодирования в каналах связи без шума и с шумом.

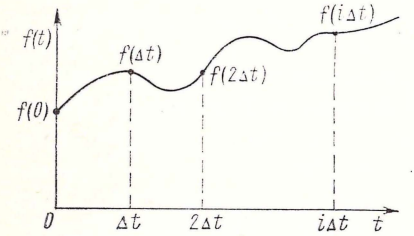


Рис. 44.

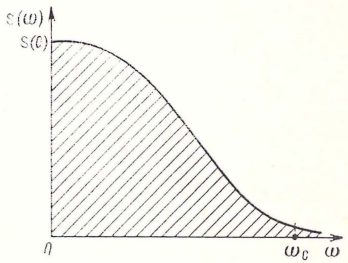


Рис. 45.

Значительная доля непрерывных сигналов, снимаемых с датчиков производства в АСУ, имеет ограниченный спектр, что обусловлено как исходными физическими измеряемыми процессами, так и ограниченной полосой пропускания устройств, входящих в датчик. К таким сигналам применима теорема Котельникова [24], в соответствии с которой функция с ограниченным спектром однозначно определяется набором своих дискретных значений, отсчитанных с временным интервалом $\Delta t = 1/2F_c = \pi/\omega_c$, где $F_c$ — ширина спектра сигнала, $\omega_c$ — частота среза функции.

Пусть задана функция $f(t)$, удовлетворяющая условиям Дирихле (рис. 44). График плотности $s(\omega)$ спектра представлен на рис. 45. Разложим функцию $s(\omega)$ в ряд, считая, что она существует в интервале от $-\omega_c$ до $\omega_c$ и период ее $T = 2\omega_c$. Тогда искомая функция

$$s(\omega) = \sum_{k=-\infty}^{\infty} D_k \exp(i2\pi k\omega/2\omega_c) = \sum_{k=-\infty}^{\infty} D_k \exp(i\pi k\omega/\omega_c),$$

где $D_k = (1/2\omega_c)\int_{-\omega_c}^{\omega_c} s(\omega)\exp(-i\pi k\omega/\omega_c)\,d\omega$.

Обозначим $\pi/\omega_c = \Delta t$. Тогда

$$s(\omega) = \sum_{k=-\infty}^{\infty} D_k \exp(ik\omega\Delta t),$$

$$D_k = (1/2\omega_c) \int_{-\omega_c}^{\omega_c} s(\omega) \exp(-ik\omega\Delta t)\, d\omega.$$

Функция $f(t)$ является непериодической и связана с плотностью амплитудного спектра интегралом Фурье:

$$f(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} s(\omega) \exp(i\omega t)\, d\omega =$$

$$= \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \sum_{k=-\infty}^{\infty} D_k \exp(ik\omega\Delta t) \exp(i\omega t)\, d\omega.$$

Выразим $f(-k\Delta t)$ через плотность амплитудного спектра $s(\omega)$:

$$f(-k\Delta t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} s(\omega) \exp[i\omega(-k\Delta t)]\, d\omega.$$

Представим $D_k$ через $f(-k\Delta t)$: $D_k = f(-k\Delta t)\,\pi/\omega_c = f(-k\Delta t)\,\Delta t$. Тогда

$$f(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \sum_{k=-\infty}^{\infty} f(-k\Delta t)\,\Delta t \exp(-ik\omega\Delta t) \exp(i\omega t)\, d\omega.$$

Окончательно получаем

$$f(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \sum_{k=-\infty}^{\infty} f(k\Delta t)\,\Delta t \exp(-ik\omega\Delta t) \exp(i\omega t) \times$$

$$\times\, dt = \sum_{k=-\infty}^{\infty} f(k\Delta t)\, \frac{\sin \omega_c (t - k\Delta t)}{\omega_c (t - k\Delta t)}.$$

Из формулы следует, что $f(t)$ может быть представлена бесконечно большим числом дискретных отсчетов $f(k\Delta t)$, умноженных на величину

$$\varphi(t) = \frac{\sin \omega_c (t - k\Delta t)}{\omega_c (t - k\Delta t)}.$$

При $k = 1$ $\varphi(t) = \dfrac{\sin \omega_c (t - \Delta t)}{\omega_c (t - \Delta t)}$ представлена на рис. 46. При $t = \Delta t$ $\varphi(\Delta t) = 1$; при $t$, кратных $k\Delta t$, $\varphi(t) = 0$. Чтобы физически реализовать эту величину, подадим бесконечно короткий единичный импульс на вход идеального фильтра, характеризующегося следующими свойствами:

1) постоянством коэффициента передачи $k(\omega)$ в полосе пропускания;

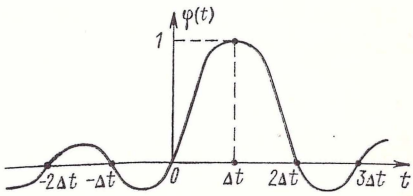


Рис. 46.

2) равенством нулю коэффициента пропускания в полосе затухания, т. е. $k(\omega) = 0$ при $\omega > \omega_c$;

3) бесконечно большой крутизной характеристики коэффициента передачи при переходе от одной полосы к другой.

Рассмотрим использование теоремы Котельникова при передаче информации. Пусть требуется передать функцию с ограниченным спектром, удовлетворяющую условиям этой теоремы. Известно $\omega_c$. Находим интервал квантования по времени $\Delta t = \pi/\omega_c$. Осуществляем выбор дискретных значений функции. Функция заменяется набором ее мгновенных значений, отсчитанных с интервалом $\Delta t$. Через линию связи передается последовательность импульсов с амплитудами $f(k\Delta t)$. Суммируем отклики от отдельных дискретных значений функции, учитывая, что в моменты $t = k\Delta t$ имеет место амплитуда лишь одного отклика. В результате получаем огибающую в виде функции $f(t)$ (рис. 47).

На рис. 48 приведена структурная схема системы передачи непрерывной информации по принципу Котельникова. Источник информации (ИИ) выдает непрерывную информацию в виде функции $f(t)$. Непрерывное напряжение проходит через ключ (К), который кратковременно открывается от импульсов с генератора (ГТИ), следующих с периодом $\Delta t = \pi/\omega_c$. С выхода К снимается последовательность отсчетов функции, т. е. $\sum_k f(k\Delta t)$, которая через линейное устройство (ЛУ), обеспечивающее согласование сигнала с линией связи, поступает в канал. На приемной стороне с ЛУ последовательность импульсов снимается на идеальный фильтр (ИФ), формирующий на выходе совокупность откликов, т. е.

$$\sum_k f(k\Delta t)\,\frac{\sin\omega_c(t-k\Delta t)}{\omega_c(t-k\Delta t)}.$$

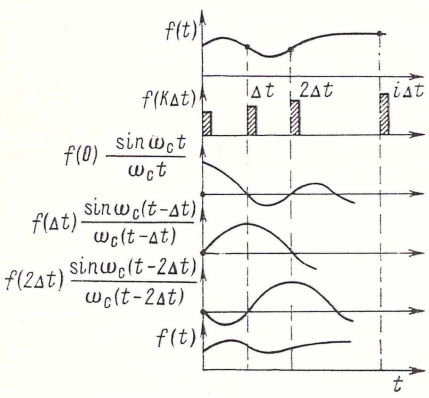


Рис. 47.

Регистрирующее устройство (РУ) выявляет принятую функцию $f(t)$.

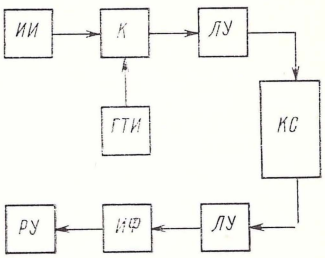


Рис. 48.

Рассмотрим реализуемость данной схемы на практике. Основные трудности, встречающиеся при реализации, следующие: 1) не удается обеспечить формирование бесконечно коротких импульсов; 2) отсутствуют каналы связи для их передачи; 3) нельзя построить идеальный фильтр.

Практически удается в определенной степени приблизиться к некоторому короткому импульсу, имея канал, удовлетворительно передающий его пятую гармонику, и фильтр с соответствующей полосой пропускания. Возникает суммарная погрешность передачи, которая делает непригодным использование этого метода в указанном виде. Тем более он непригоден в канале с помехами. Если погрешность является допустимой, то необходимо исключить влияние помех, которые действуют в канале связи. Для этого используют квантование по амплитуде.

Введем детерминированную сетку квантования по амплитуде. Обозначим шаг квантования через $\delta$. На приемной стороне будем считать разрешенными лишь заданные уровни квантования. На рис. 49 отмечены уровни квантования функции. По теореме Котельникова вводим квантование функции по времени и в линию связи передаем импульсы, кратные $\delta$, т. е. заранее вносим погрешность на передающей стороне, не превышающую по амплитуде $\xi_i$. Процесс квантования по амплитуде дает погрешность $\xi_i = f(i\Delta t) - k\delta$, где $k\delta$ выбирается ближайшим к соответствующему значению функции $f(t)$.

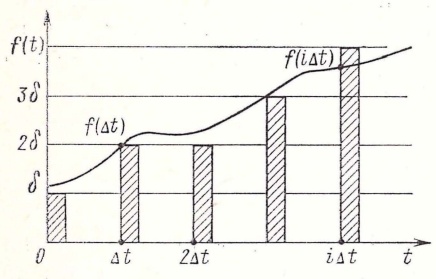


Рис. 49.

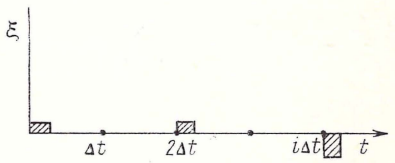


Рис. 50.

На рис. 50 показан график погрешности $\xi_i$. Целесообразно стремиться, чтобы усредненное значение $\xi_i$ по многим реализациям было равно нулю. Если это не получается, то необходимо изменить направление аппроксимации для точек функции, попавших в середину шага квантования. Последовательность импульсов поступает в линию связи. Если в линии существуют помехи, то на приемной стороне имеем значение функции $f(i\Delta t) = k\delta + \psi_i$, где $\psi_i$ — значение помехи $i$-го отсчета. На полученную последовательность импульсов накладываем известную нам шкалу квантования и аппроксимируем импульсы по данной шкале. При такой аппроксимации устра-

няется погрешность, связанная с действием помехи в линии связи, если $\psi_i < \delta/2$.

В действительности значение $\psi_i$ зависит от протяженности линии связи. Для снижения $\psi_i$ следует уменьшать расстояние между передающими и приемными станциями с помощью ретрансляционных станций.

Недостатком рассмотренного способа передачи является использование амплитудного сигнального признака, обладающего низкой помехоустойчивостью. Поэтому при передаче на большие расстояния более полезно представление каждого квантованного уровня кодом, обеспечивающим заданную помехоустойчивость передачи. При этом получаем передачу с использованием кодоимпульсной модуляции.

Пусть имеется $M$ уровней квантования по амплитуде, т. е. $f_{\max} - f_{\min} = (M-1)\delta$. Каждый уровень представим комбинацией избыточного кода, т. е. будем считать его сообщением. По известной статистике ошибок в канале связи и по заданной вероятности ошибки $[P_0]_{\text{доп}}$ определяем необходимые обнаруживающую и исправляющую способности кода. Строим код и каждый уровень амплитуды передаем соответствующей кодовой комбинацией.

Пример. Передача непрерывной функции с помощью кодоимпульсной модуляции при использовании двоичного кода.

Положим $f_{\max} - f_{\min} = 6\delta$, т. е. $M = 7$. Пусть из статистических характеристик канала связи получено, что код должен обнаруживать одну ошибку, т. е. $r = 1$, $d = 2$. Так как $M = 2^{n-1}$, то $n = 4$, $m = 3$. На информационных позициях строим неизбыточный двоичный код. Получаем кодовые комбинации 0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111. При этом значение символа на контрольной позиции $x_4 = \sum_{i=1}^{3} x_3$. Код вида 0000 не используем, остальные соответствуют принятым уровням квантования по амплитуде.

## § 2. Энтропия непрерывного дискретизированного процесса

Пусть существует некоторая непрерывная функция $f(t)$, которая дискретизирована по времени. Известна статистика этих отсчетов, т. е. проанализирован данный процесс в течение значительных интервалов времени и установлены вероятности появления любого уровня $x(t)$. Построим плотность распределения вероятности $W(x)$ (рис. 51). Между отсчетами имеется некоторый шаг квантования по амплитуде $\Delta x$. Рассмотрим случай, когда $\Delta x = \text{const}$. Определим вероятность появления некоторого отсчета $x_i$. Отметим вок-

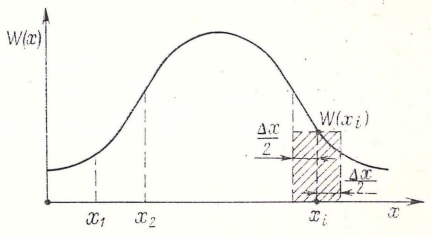


Рис. 51.

руг этого отсчета зону $\Delta x$. Площадь этой зоны есть вероятность

$$p(x_i) = \int_{x_i-\Delta x/2}^{x_i+\Delta x/2} W(x)\,dx.$$

Заменим полученную площадь прямоугольником, тогда $p(x_i) = W(x_i)\Delta x$. Энтропией отсчета назовем меру неопределенности отсчета данного уровня, которая снимается при появлении этого отсчета, т. е. количество информации, передаваемое данным отсчетом функции, и обозначим через $H(x_i)$. Тогда

$$H(x_i) = -p(x_i)\log_2 p(x_i) = -W(x_i)\Delta x \log_2 [W(x_i)\Delta x] =$$

$$= -W(x_i)\Delta x \log_2 W(x_i) - W(x_i)\Delta x \log_2 \Delta x.$$

Так как

$$\lim_{\Delta x\to 0} H(x_i) = H_i(x) = -\lim_{\Delta x\to 0} W(x_i)\Delta x \log_2 W(x_i) -$$

$$-\lim_{\Delta x\to 0} W(x_i)\Delta x \log_2 \Delta x,$$

то для всей совокупности отсчетов

$$H(x) = -\sum_{-\infty}^{\infty} \lim_{\Delta x\to 0} W(x_i)\Delta x \log_2 W(x_i) -$$

$$-\sum_{-\infty}^{\infty} \lim_{\Delta x\to 0} W(x_i)\Delta x \log_2 \Delta x =$$

$$= -\int_{-\infty}^{\infty} W(x)\log_2 W(x)\,dx) - \int_{-\infty}^{\infty} W(x)\log_2 \Delta x dx.$$

Второй член не определен и поэтому обычно не рассчитывается, т. е.

$$H_{\text{п}}(x) = -\int_{-\infty}^{\infty} W(x)\log_2 W(x)\,dx,$$

где $H_{\text{п}}(x)$ — дифференциальная, или приведенная, энтропия. Из полученных соотношений следует:

1) дифференциальная энтропия определяется статистикой отсчетов функции. Как ранее показано, наибольшее значение дифференциальной энтропии соответствует нормальному закону распределения отсчетов функции;

2) энтропия зависит от амплитуды функции и выбора шага квантования.

Различают системы с равномерным и неравномерным шагом квантования.

*Система с равномерным шагом квантования.* Пусть имеется некоторая функция $x(t)$ (рис. 52,а), которая характери-

зуется тем, что $P(x)$ (рис. 52,б) представляется равномерным законом распределения, т. е. значения амплитуд функции равновероятны. В этом случае согласно дифференциальной энтропии $H_{\text{п}}(x)$ целесообразно иметь равномерный шаг квантования по амплитуде. Тогда каждый отсчет несет одинаковое количество информации и можно передавать любой отсчет в виде импульса заданной амплитуды, кратной $\delta$, либо равномерным кодом с требуемой избыточностью.

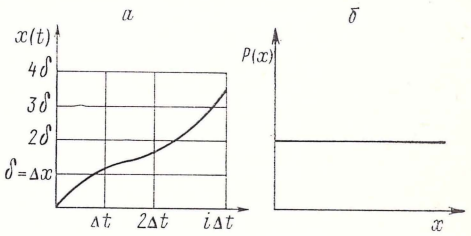

Рис. 52.

*Система с неравномерным шагом квантования.* В этом случае передаваемая функция $x(t)$ имеет распределение $W(x)$, отличное от равномерного (рис. 53,а,б). Наиболее вероятное значение есть математическое ожидание $M(x)$. Целесообразно использовать неравномерный шаг квантования с уменьшением $\Delta x$ при приближении к точке $M(x)$, т. е. $\Delta x_2 > \Delta x_1$. Тогда наиболее вероятные значения будут передаваться с большей точностью. Любой отсчет будет переносить одинаковое количество информации. Передача возможна либо импульсами определенной амплитуды, либо равномерным кодом. Для повышения количества передаваемой информации необходимо перейти к передаче отклонений функции от математического ожидания.

*Система с передачей отклонений функции от математического ожидания.* Пусть имеется произвольная непрерывная функция $x(t)$ и известен закон распределения $W(x)$. Вместо значений $x(t)$ будем передавать $x(t) - M(x)$, т. е. отклонения функции от математического ожидания. Если отклонения невелики, то можно использовать нормальный закон распределения для описания процессов с нулевым математическим ожиданием. Тогда, как показано ранее, приведенная энтропия

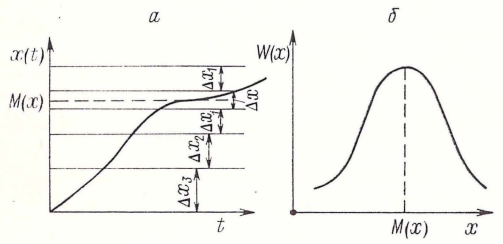

Рис. 53.

$$H_{\text{п}}(x) = H_{\max} = \log_2 \sqrt{2\pi e \sigma^2},$$

где $\sigma$ — дисперсия процесса, $\sigma^2 = P_{\text{с}}$.

Таким образом, при данном способе передачи получаем максимальную дифференциальную энтропию. Отсюда можно сделать следующие выводы: 1) для повышения дифференциальной энтропии следует передавать вместо функции ее отклонения от математического ожидания; 2) наибольшее количество информации передается в случае, если в канале связи действует помеха с нормальным законом распределения.

## § 3. Скорость передачи непрерывной дискретизированной информации и пропускная способность канала связи

Пусть в канал связи с шумом поступает последовательность отсчетов $X$, на выходе канала принимаем множество отсчетов $Y$, искаженных помехами. Количество информации, передаваемое через канал связи: $I(X, Y) = H(Y) - H(Y/X)$, где $H(Y)$ — энтропия выходного множества отсчетов, $H(Y/X)$ — условная энтропия, учитывающая действие помех в линии связи.

Так как непрерывная функция передается в виде дискретных отсчетов, то воспользуемся понятием дифференциальной энтропии. Пусть известны плотность вероятности распределения отсчетов и помехи $W(x+\xi) = W(y)$, плотность вероятности распределения амплитуд отсчетов на входе канала связи $W(x)$ и плотность вероятности распределения амплитуд помех в канале связи $W(\xi)$. Скорость передачи информации будем измерять количеством информации, передаваемым в одном отсчете. Тогда

$$H(Y) = -\int_{-\infty}^{\infty} W(x+\xi) \log_2 W(x+\xi)\, d(x+\xi).$$

Условная энтропия $H(Y/X)$ есть потери информации, или количество ложной информации, вносимой помехой в передаваемый сигнал:

$$H(Y/X) = H(\xi) = -\int_{-\infty}^{\infty} W(\xi) \log_2 W(\xi)\, d\xi,$$

где $H(\xi)$ — энтропия помехи. Скорость передачи информации определится следующим образом:

$$R = I(X, Y) = -\int_{-\infty}^{\infty} W(x+\xi) \log_2 (x+\xi)\, d(x+\xi) + \\ + \int_{-\infty}^{\infty} W(\xi) \log_2 W(\xi)\, d\xi.$$

Максимальная скорость обеспечивается, если статистика отсчетов совпадает со статистикой помехи. Для канала с нормальным распределением помехи

$$C = R_{\max} = -\int_{-\infty}^{\infty} W(x+\xi) \log_2 (x+\xi)\, d(x+\xi) + \log_2 \sqrt{2\pi e \sigma^2}.$$

Рассмотрим зависимость пропускной способности от физических характеристик канала связи. Пусть статистика отсчетов и помехи одинакова. Повысим пропускную способность за счет расширения полосы пропускания канала связи. Пропускная способность канала связи

$$C = F_к \log_2 (1 + P_с/P_п),$$

где $F_к$ — ширина полосы пропускания канала связи.

Введем понятие удельной мощности помехи $P_{п_0}$. Удельная мощность помехи есть мощность помехи, приведенная к единице полосы пропускания, т. е. $P_{п_0} = P_п/F_к$. Тогда

$$C = F_к \log_2 (1 + P_с/P_{п_0}F_к).$$

Предел пропускной способности при обеспеченном расширении полосы

$$\lim_{F_к \to \infty} C = \lim F_к \log_2 (1 + P_с/P_{п_0}F_к) = (P_с/P_{п_0}F_к) \log_2 e = (P_с/P_п)\log_2 e,$$

так как $\lim (P_с/P_{п_0}F_к) = 0$ при $F_к \to \infty$.

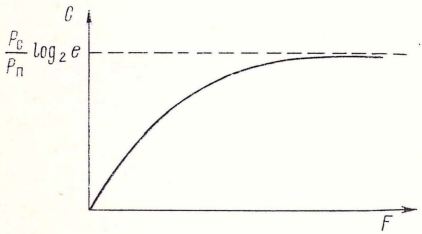


Рис. 54.

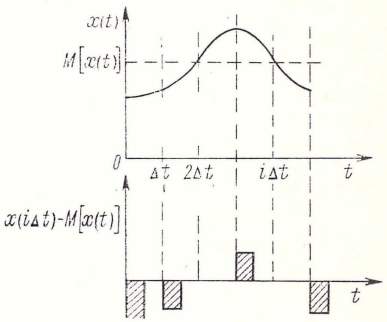


Рис. 55.

На рис. 54 представлена зависимость $C = \varphi(F_к)$, откуда следует, что при любом способе передачи и идеально широком канале связи пропускная способность ограничена отношением мощностей сигнала и помехи. Целесообразно выбирать канал таким, чтобы $P_п$ была минимальной.

Положим, что передается некоторая функция $x(t)$ и известно ее математическое ожидание $M[x(t)]$. Для повышения скорости передачи информации будем передавать отклонения функции от математического ожидания в виде некоторого амплитудно-полярного сигнала (рис. 55). Определим пропускную способность канала связи для такого сигнала. Пусть в канале связи действует помеха с нормальным законом распределения амплитуд. На приемной стороне установлены пороги срабатывания $X$, $-X$. Вероятность ошибки определим как $p = p(0)\,p(1/0) + p(1)\,p(0/1)$, где $p(0)$, $p(1)$ — вероятности передачи соответствующих значений сигнального признака.

Ввиду малого значения амплитуд передаваемых сигналов будем считать, что распределение $W(\xi)$ при наличии сигнала смещается незначительно. Тогда

$$p(1/0) = \int_X^\infty W(\xi)\, d\xi, \quad p(0/1) = \int_{-\infty}^{-X} W(\xi)\, d\xi.$$

Положим, что отклонения равновероятны, т. е. $p(0) = p(1) = 1/2$. Тогда вероятность искажения одного отсчета функции

$$p = \frac{1}{2}\left[\int_X^\infty W(\xi)\, d\xi + \int_{-\infty}^{-X} W(\xi)\, d\xi\right] = \frac{1}{2}\left[1 - \int_{-X}^{X} W(\xi)\, d\xi\right] =$$

$$= \frac{1}{2}\left[1 - 2\int_0^X W(\xi)\, d\xi\right] = \frac{1}{2} - \int_0^X W(\xi)\, d\xi.$$

Учитывая, что $W(\xi) = \left(1/\sqrt{2\pi\sigma^2}\right) \exp\left(-\xi^2/2\sigma^2\right)$, получаем

$$p = \frac{1}{2} - \int_0^X \left(1/\sqrt{2\pi\sigma^2}\right) \exp\left(-\xi^2/2\sigma^2\right) d\xi = \frac{1}{2} - \Phi\left(X, \xi^2/2\sigma^2\right),$$

где $\Phi$ — функция Лапласа, значение которой находим по таблицам. Так как $X$ определяется амплитудой передаваемого сигнала, т. е. минимальным отклонением, то получаем, что $p$ зависит от отношения мощности сигнала к мощности помехи. Зависимость $p = \varphi(P_c/P_п)$ представлена на рис. 56. Видно, что вероятность искажения определяется уровнем помех в канале связи и необходимой точностью воспроизведения функции. Зная $p$, определим пропускную способность $C = 1 + p\log_2 p + (1-p)\log_2(1-p)$.

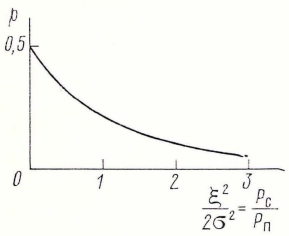


Рис. 56.

Анализ показывает, что пропускная способность в данном случае оказывается близкой к пропускной способности непрерывного канала. Поэтому в данном варианте передачи обеспечиваются наибольшая пропускная способность и наибольшая скорость передачи информации.

## § 4. Погрешности квантования непрерывной функции по времени и по амплитуде

При квантовании функции по методу Котельникова возникает погрешность квантования по времени и по амплитуде. Погрешность квантования обусловливается следующими причинами.

1. Спектр реально существующих функций, как правило, не является конечным. При передаче с последующей дискретизацией отсекаются гармонические составляющие выше $\omega_c$

и имеются потери энергии, определяемые коэффициентом $\nu$:

$$\nu = E_0/E_c = \int_{\omega_c}^{\infty} s^2(\omega)\, d\omega / \int_0^{\omega_c} s^2(\omega)\, d\omega,$$

где $E_0$ — энергия ошибки; $E_c$ — энергия сигнала. Для уменьшения погрешности необходимо так выбрать $\omega_c$, чтобы $\upsilon$ был в заданных пределах, но в любом случае остается погрешность, связанная с невозможностью передачи всего спектра функции $x(t)$.

2. Не удается воспроизвести отклик идеального фильтра в интервале времени до появления входного импульса. Пусть имеется некоторый отсчет функции, который поступает на идеальный фильтр. Тогда по теореме Котельникова получаем отклик, представленный на рис. 57. Сигнал с выхода фильтра возникает раньше, чем поступит на вход. В интервале $T$ реальный фильтр не может воспроизвести сигнал с предсказанием, поэтому возникает погрешность, которую можно несколько уменьшить, если увеличить частоту отсчетов.

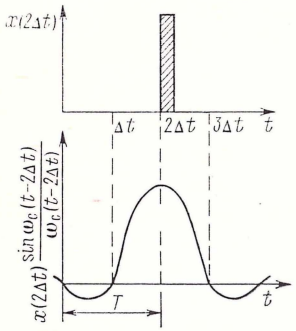


Рис. 57.

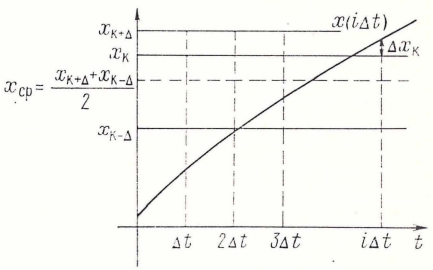


Рис. 58.

Названные причины не позволяют проводить квантование с периодом, вычисляемым по теореме Котельникова. При передаче периодических сигналов приходится увеличить частоту квантования в 2 — 3 раза.

Оценим погрешность квантования по амплитуде. Пусть осуществляется передача непрерывной информации с квантованием по времени и по амплитуде. Оценим погрешность, связанную с отклонением действительного значения функции от квантованного значения, и определим оптимальное значение шага квантования по амплитуде. На рис. 58 дана функция $x(t)$, квантованная по амплитуде и по времени. Пусть в момент $i\Delta t$ вместо значений функции $x(i\Delta t)$ передается квантованное значение $x_k$. Оценим погрешность. Положим, что шаг квантования $x_{k+\Delta} - x_{k-\Delta}$ является переменным и уровень расположен произвольно в зоне шага квантования. Введем понятие отклонения функции от уровня $x_k$: $\Delta x_k = x(i\Delta t) - x_k$.

Найдем математическое ожидание, т. е. среднее отклонение действительного значения функции от квантованного уровня:

$$M(\Delta x) = \sum_{k=1}^{M} \Delta x_k p(x_k),$$

где $M$ — максимальное число уровней при квантовании. Считаем, что различные уровни функции являются равновероятными, и определим среднее отклонение функции от уровня $x_k$:

$$M(\Delta x_k) = \int_{x_{k-\Delta}}^{x_{k+\Delta}} (x - x_k)\, p(x)\, dx,$$

где $p(x)$ — вероятность появления амплитуды функции $x$, $x - x_k$ — отклонение функции от квантованного по амплитуде значения.

Допустим, что вероятность $p(x)$ в зоне уровня $x_k$ — величина постоянная и равная некоторому значению $p(x_{ср})$, т. е. $p(x) = p(x_{ср})$, где $x_{ср} = (x_{k+\Delta} + x_{k-\Delta})/2$. Тогда (см. [18]):

$$M(\Delta x_k) = p(x_{ср}) \int_{x_{k-\Delta}}^{x_{k+\Delta}} (x - x_k)\, dx =$$

$$= [p(x_{ср})/2]\, (x - x_k)^2 \Big|_{x_{k-\Delta}}^{x_{k+\Delta}} =$$

$$= [p(x_{ср})/2]\, [(x_{k+\Delta} - x_k)^2 - (x_{k-\Delta} - x_k)^2].$$

Усредненное отклонение $\Delta x_k$ определяется вероятностью возникновения функции в заданном уровне и амплитудой шага квантования. Оценим дисперсию отклонения

$$D(\Delta x_k) = \int_{x_{k-\Delta}}^{x_{k+\Delta}} (x - x_k)^2 p(x)\, dx =$$

$$= p(x_{ср}) \int_{x_{k+\Delta}}^{x_{k+\Delta}} (x - x_k)^2\, dx =$$

$$= [p(x_{ср})/3]\, [(x_{k+\Delta} - x_k)^3 - (x_{k+\Delta} - x_k)^3].$$

Найдем связь между $x_k$, $x_{k+\Delta}$ и $x_{k-\Delta}$. Выявим минимум дисперсии. Для этого возьмем производную

$$\partial D(\Delta x_k)/\partial x_k = p(x_{ср})\, [-(x_{k+\Delta} - x_k)^2 + (x_{k+\Delta} - x_k)^2] = 0.$$

Так как $p(x_{ср}) \neq 0$, то $(x_{k-\Delta} - x_k)^2 = (x_{k+\Delta} - x_k)^2$, или $\pm (x_{k-\Delta} - x_k) = \pm (x_{k+\Delta} - x_k)$. Решая это уравнение, получаем:

а) $x_{k+\Delta} - x_k = (x_{k-\Delta} - x_k)$, $x_{k+\Delta} = x_{k-\Delta}$, т. е. зоны квантования нет; следовательно, это решение не удовлетворяет начальным условиям;

б) $x_{k+\Delta} - x_k = -(x_{k-\Delta} - x_k)$, $x_k = (x_{k+\Delta} + x_{k-\Delta})/2$, т. е. значение уровня квантования должно выбираться в середине

шага квантования. Подставляя решение в математическое ожидание и дисперсию, получаем

$$M(\Delta x_k) = [p(x_{ср})/2]\,[(x_{k+\Delta} - x_k)^2 - (x_{k-\Delta} - x_k)^2] = \\ = [p(x_{ср})/2]\,[(\delta_k/2)^2 - (\delta_k/2)^2] = 0,$$

$$D(\Delta x_k) = [p(x_{ср})/3]\,[(x_{k+\Delta} - x_k)^3 - (x_{k-\Delta} - x_k^2)] = \\ = p(x_{ср})\,\delta_k^3/12,$$

где $x_{k+\Delta} - x_k = \delta_k$ и представляет собой шаг квантования уровня $k$.

Значение дисперсии зависит от статистики отсчетов функции, и поэтому имеется возможность построить оптимальную шкалу квантования. При условии, что значения $x_{ср\,k}$ являются некоррелированными, суммарная дисперсия по всем уровням квантования составит

$$D(\Delta x) = \sum_{k=1}^{M} [p(x_{срk})\,\delta_k^3/12].$$

Если шаг квантования является постоянным, т. е. $\delta_k = \delta$, то

$$D = \sum_{k=1}^{M} [p(x_{срk})\,\delta^3/12] = \delta^2 \sum_{k=1}^{M} [p(x_{ср\,k})\,\delta/12].$$

Рассмотрим сумму $\sum_{k=1}^{M} p(x_{ср\,k})\,\delta$. Это есть вероятность попадания функции $x(t)$ в один из $M$ уровней, т. е. охватывает полную группу событий, а поэтому сумма равна единице. Окончательно получаем $D = \delta^2/12$.

Таким образом, при переходе к равномерному шагу квантования дисперсия в меньшей степени зависит от $\delta$, т. е. можно расширить шаг квантования при сохранении постоянства дисперсии. Равномерный шаг квантования можно использовать в двух случаях: 1) все значения амплитуд функций равновероятны; 2) значения амплитуд функций неравновероятны, но каждый уровень кодируется неравномерным кодом в соответствии с принципами оптимального кодирования дискретной информации.

## § 5. Оптимальное квантование непрерывной функции по амплитуде

Пусть имеется функция $x(t)$ и известен закон распределения ее амплитудных значений $p(x)$. Требуется выбрать такой неравномерный шаг квантования, чтобы обеспечить минимум дисперсии

$$D = \sum_{k=1}^{M} [p(x_k)\,\delta_k^3/12],$$

где $p(x_k)$ — вероятность попадания амплитуды функции в интервал $\delta_k$. Введем переменную $z_k = [p(x_k)]^{1/3}\,\delta_k$. Тогда

$D=\sum_{k=1}^{M}(z_k^3/12)$. Чтобы обеспечить минимум дисперсии, необходимо выполнить условие $\sum_{k=1}^{M} p(x_k)\,\delta_k = 1$, отсюда $\sum_{k=1}^{M} z_k = s$, где $s$ — некоторое постоянное число.

Требуется найти такое $z_k$, для которого дисперсия минимальна. Для этого воспользуемся методом вариационного исчисления. Введем функцию $\varphi = s - \sum_{k=1}^{M} z_k$, тогда $f_1 = D + \lambda\varphi$, где $\lambda$ — неопределенный множитель:

$$\partial f_1/\partial z_k = \partial D/\partial z_k + \lambda\partial\varphi/dz_k.$$

Пусть $z_k$ фиксировано, в таком случае $\partial f_1/\partial z_k = z_k^2/4 - \lambda = 0$, $z_k = 2\sqrt{\lambda}$, что соответствует минимуму дисперсии. Выразим $D_{\min}$ через $s$. Так как все $z_k$ одинаковы, то можно записать: $Mz_k = s$, $z_k = s/M$. Следовательно, $D = D_{\min} = Ms^3/12M^3 = s^3/12M^2$.

Можно показать, что данная величина дисперсии оказывается меньше в 5 — 10 раз дисперсии для равномерного шага квантования.

*Построение оптимальной шкалы квантования.* Если функция $x(t)$ меняется от 0 до $x_m$ и квантуется с неравномерным шагом квантования, то некоторый $i$-й уровень функции может быть записан как $x_i = \sum_{j=0}^{i} \delta_j$, причем $[p(x_j)]^{1/3}\delta_j = z_k = s/M$, $\delta_j = s/M\,[p(x_j)]^{1/3}$.

Считая, что число уровней растет неограниченно, т. е. $M \to \infty$, введем новую переменную $y$, для которой $1/M \to dy$. Получаем функцию

$$x(y) = \int_0^y [Q/[p(y)]^{1/3}]\,dy.$$

Путем равномерного разбиения $y$ в интервале от 0 до $x_m$ можно получить неравномерное разбиение $x$, что соответствует оптимальному квантованию по уровню. Определим значение $Q$. Очевидно, что

$$x_m = \int_0^{x_m} [Q/[p(y)]^{1/3}]\,dy = Qs_1, \text{ отсюда}$$

$$s_1 = \int_0^{x_m} [p(y)]^{-1/3}\,dy,\ Q = x_m/s_1.$$

Окончательно получаем формулы, определяющие оптимальную шкалу квантования по уровню (см. [18]):

$$x(y) = \int_0^y (x_m/s_1)\,[p(y)]^{-1/3}\,dy,\ s_1 = \int_0^{x_m} [p(y)]^{-1/3}\,dy.$$

Рассмотрим принцип построения оптимальной шкалы квантования. Пусть известно распределение $p(x)$ и $0 \leqslant x(t) \leqslant x_m$.

Разобьем интервал от 0 до $x_m$ по оси $y$ (рис. 59) на равные части. Для точки $y_1$ находим

$$x_1 = \int_0^{y_1} (x_m/s_1)\,[p\,(y)]^{-1/3}\,dy, \text{ где } s_1 = \int_0^{x_m} [p\,(y)]^{-1/3}\,dy.$$

Повторяем эту операцию для каждой точки $y_i$ и строим зависимость $x\,(y)$. Полученные значения $x$ есть уровни квантования функции. Далее относительно каждого уровня выбирается шаг квантования по вышеприведенной методике и уровни располагаются посредине шагов квантования.

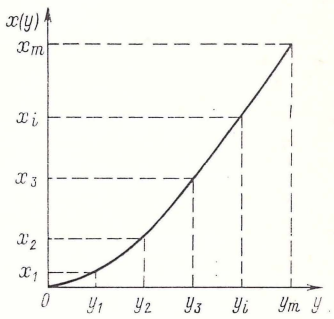


Рис. 59.

Пример 1. Оценка дисперсии и принцип построения оптимальной шкалы квантования по уровню для функции, подчиняющейся нормальному закону распределения амплитуд отсчетов.

Пусть задана функция $x\,(t)$ и амплитуды отсчетов подчинены нормальному закону распределения

$$p\,(x) = \left(1/\sqrt{2\pi\sigma_x^2}\right) \exp\left(-x^2/2\sigma_x^2\right),$$

где $\sigma_x$ — эффективное значение функции.

Так как $\sum_{k=1}^{M} p\,(x_k)\,\delta_k = 1$, $\sum_{k=1}^{M} z_k = s$, где $z_k = [p\,(x)]^{1/3}\,\delta_k$, то $\sum_{k=1}^{M} [p\,(x)]^{1/3}\,\delta_k = s$. При изменении $k$ от $-\infty$ до $\infty$ и конечной амплитуде функции $\delta_k$ переходит в $dx$. Тогда для рассматриваемого примера получаем

$$s = \int_{-x_m}^{x_m} \left(1/\sqrt{2\pi\sigma_x^2}\right)^{1/3} \exp\left(-x^2/6\sigma_x^2\right) dx =$$

$$= 2\int_0^{x_m} \left(1/\sqrt{2\pi\sigma_x^2}\right)^{1/3} \exp\left(-x^2/6\sigma_x^2\right) dx =$$

$$= \sqrt{3}\,(\sigma_x\sqrt{2\pi})^{2/3}\,\Phi\,(x_m/\sigma_x\sqrt{3}),$$

где $\Phi\,(x_m/\sigma_x\sqrt{3}) = (2/\sqrt{\pi})\int_0^{x_m} \exp(-t^2/2)\,dt$, $t = x/\sigma_x\sqrt{3}$.

Окончательно

$$D_{\min} = s^3/12M^2 = (\sqrt{3})^3\,(\sigma_x\sqrt{2\pi})^2\,\Phi\,(x_m/\sigma_x\sqrt{3})^3/12M^2.$$

Если сравнить полученный результат с равномерным шагом квантования, то имеем снижение погрешности $\upsilon = D_{\min}/D$. На рис. 60 представлена зависимость $\upsilon = \varphi(x_m/\sigma_x)$. Так как $x_m$ обычно больше, чем $\sigma_x$, то всегда $D_{\min} < D$.

Пример 2. Оптимальное квантование функции, подчиняющейся двустороннему экспоненциальному распределению отсчетов по амплитуде

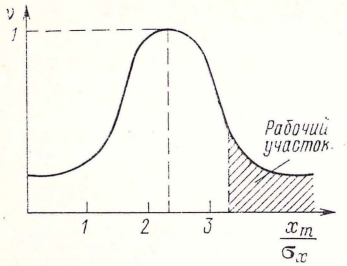


Рис. 60.

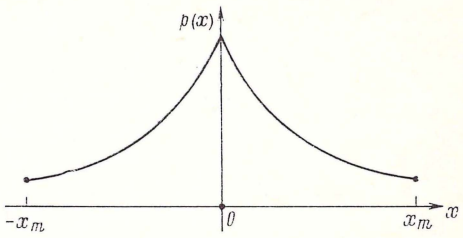


Рис. 61.

На рис. 61 представлена зависимость $p(x) = \frac{a}{2} \exp(-a|x|)$, в этом случае

$$s = \int_{-x_m}^{x_m} (a/2)^{1/3} \exp(-a|x|/3)\, dx = 6\,(a/2)^{1/3}[1 - \exp(-ax_m/3)]$$

и $D_{\min} = 9\,[1 - \exp(-ax_m/3)]^3$.

Для оптимального квантования $D_{\min}$ всегда будет меньше $D$, соответствующей равномерному шагу квантования.

# *Глава VI*

## ВЫДЕЛЕНИЕ СИГНАЛА НА ФОНЕ ПОМЕХ

### § 1. Геометрическое представление сигнала в пространстве

Каждый элемент сигнала может быть представлен точкой в пространстве с соответствующей системой координат, размерность которой равна числу используемых сигналов. Кодовое расстояние $d$ определяется расстоянием между точками. Рассмотрим, как нужно выбирать сигналы, чтобы $d_{\min}$ было наибольшим.

На рис. 62 представлена двумерная модель для двух сигналов. Определим $d$ через проекции векторов

$$d = \sqrt{[(x_{21} - x_{11})^2 - (x_{12} - x_{22})^2]} .$$

Если отображается $n$ сигналов, то кодовое расстояние между сигналами $x_1(t)$ и $x_2(t)$

$$d = \sqrt{\sum_{j=1}^{n} (x_{2j} - x_{1j})^2} ,$$

где $x_{1j}$, $x_{2j}$ — проекции векторов $x_1$ и $x_2$ на ось $j$.

Перейдем к физическим характеристикам сигнала. Рассмотрим передачу непрерывного сигнала $x(t)$, дискретизируемого согласно теореме Котельникова.

Пусть длительность сигнала $x(t)$ равна $T$, тогда число отсчетов $n = T/\Delta t$, где $\Delta t = 1/2F$, $F$ — ширина спектра. Получаем $2FT$ сигналов, которые различны, если функция непрерывно меняется. Найдем энергию сигнала для $j$-го отсчета. В соответствии

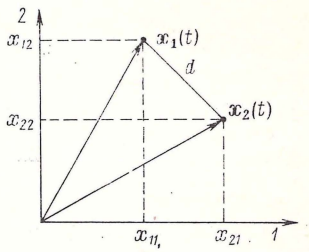


Рис. 62.

с теоремой Котельникова на выходе идеального фильтра от импульса $j$-го отсчета возникает отклик вида

$$x_j(t) = x_j \sin[\omega_c(t - j\Delta t)]/[\omega_c(t - j\Delta t)].$$

Энергия этого сигнала

$$E_{cj} = \int_{-\infty}^{\infty} \{x_j^2 \sin^2[\omega_c(t - j\Delta t)]/[\omega_c(t - j\Delta t)]^2\}\, dt,$$

энергия сигнала для $n$ отсчетов $E_c = \sum_{j=1}^{n} E_{cj}$. Учитывая, что $\omega_c = 2\pi F$, и обозначая $\omega_c(t - j\Delta t) = 2\pi F(t - j\Delta t) = \beta$, получаем

$$E_c = \sum_{j=1}^{n} (x_j^2/2\pi F) \int_{-\infty}^{\infty} (\sin^2\beta/\beta^2)\, d\beta.$$

Ввиду того, что $\int_{-\infty}^{\infty} (\sin^2\beta/\beta^2)\, d\beta = \pi$, находим

$$E_c = (1/2F) \sum_{j=1}^{n} x_j^2, \text{ или } \sum_{j=1}^{n} x_j^2 = 2FE_c.$$

Так как $n$ — это число сигналов или отсчетов, то представим, что $x_j$ есть некоторая составляющая сигнала $x(t)$. Тогда получаем $n$-мерное пространство, в котором $x_j$ — обобщенная проекция сигнала $x$ на ось $j$:

$$|x| = \sqrt{\sum_{j=1}^{n} x_j^2} = \sqrt{2FE_c}.$$

Учитывая, что $E_c = P_c T$, получаем $|x| = \sqrt{2FP_cT}$. Данное выражение справедливо для дискретного или непрерывного сигнала $x(t)$, обладающего мощностью $P_c$ и шириной спектра $F$, причем сигнал тем эффективнее, чем больше $F$.

Пусть в пространстве имеют место обобщенные сигналы $x_j(t)$ и $x_i(t)$ и каждый из них разлагается в соответствии с теоремой Котельникова. Тогда получаем $n$ отсчетов для каждого сигнала, т. е. $n$ осей координат. Если сигналы дискретны, то расстояние между ними

$$d = \sqrt{\sum_{k=1}^{2FT} (x_{jk} - x_{ik})^2},$$

где $x_{jk}$, $x_{ik}$ — проекции векторов сигналов на ось $k$. Если сигналы $x_j(t)$ и $x_i(t)$ непрерывны во времени, то

$$d_{ij} = \sqrt{\int_0^T [x_j(t) - x_i(t)]^2\, dt} =$$

$$= \sqrt{\int_0^T [x_j^2(t) - 2x_j(t)\, x_i(t) + x_i^2(t)]\, dt}.$$

Если амплитуды сигналов одинаковы, то $x_j(t) = x_i(t)$, а отсюда

$$d_{ij} = \sqrt{2\int_0^T x_j^2(t)\, dt - 2\int_0^T x_j(t)\, x_i(t)\, dt} = \sqrt{2(E_c - \lambda_{ij})},$$

где $E_c$ — энергия сигнала, $\lambda_{ij}$ — коэффициент корреляции. Чтобы увеличить расстояние $d_{ij}$, нужно иметь отрицательный коэффициент $\lambda_{ij}$.

Пусть задано $n$ сигналов и расстояние между ними — наибольшее. Ввиду того что усредненное кодовое расстояние ориентировано на наименьшее, расстояния между сигналами выгодно иметь одинаковыми. Когда сигналы исходят из точки в пространстве, то расстояния одинаковы, если векторы упираются в вершины правильного многоугольника, а начала векторов находятся в центре тяжести, т. е. $\sum_{j=1}^{n} \mathbf{x}_j = 0$. Определим квадрат суммы векторов:

$$(\mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_j + \ldots + \mathbf{x}_n)^2 = \\ = \sum_{j=1}^{n} \mathbf{x}_i^2 + \sum_{j=1}^{n} \sum_{i=1}^{n-1} \mathbf{x}_i \mathbf{x}_j = 0.$$

Ранее получено, что $|x\,i| = \sqrt{2FE_c}$, а поэтому

$$\sum_{i=1}^{n} \mathbf{x}_i^2 = 2FE_c n, \quad \sum_{j=1}^{n} \sum_{i=1}^{n-1} \mathbf{x}_i \mathbf{x}_j = n(n-1)\, 2E_c F \cos\varphi.$$

Находим угол $\varphi$ между векторами из уравнения $2E_c Fn + n(n-1)\,2E_c F\cos\varphi = 0$, или $\cos\varphi = -1/n - 1$, зная который получаем коэффициент корреляции

$$\lambda_{ij} = \int_0^T x_i(t)\, x_j(t)\, dt.$$

Представим каждый сигнал с помощью теоремы Котельникова, тогда

$$\lambda_{ij} = \int_0^T \left[ \sum_{k=1}^{n} x_{ik} \frac{\sin^2 2\pi F(t - k\Delta t)}{2\pi F(t - k\Delta t)} \times \right. \\ \left. \times \sum_{l=1}^{n} x_{jl} \frac{\sin 2\pi F(t - l\Delta t)}{2\pi F(t - l\Delta t)} \right] dt,$$

где $k$ — номер отсчета сигнала $x_i(t)$, $l$ — номер отсчета сигнала $x_j(t)$. Так как

$$\int_0^T \frac{\sin 2\pi F(t - k\Delta t) \sin 2\pi F(t - l\Delta t)}{2\pi F(t - k\Delta t)\, 2\pi F(t - l\Delta t)} = \begin{cases} 0, & k \neq l, \\ 1/2F, & k = l, \end{cases}$$

то $\lambda_{ij} = (1/2F) \sum_{k=1}^{n} x_{jk} x_{ik} = (1/2F)\, \mathbf{x}_j \mathbf{x}_l = (1/2F)\, |x_j|\, |x_l| \cos\varphi =$

$$= (1/2F) \sqrt{2FE_c}\, \sqrt{2FE_c} \cos\varphi = E_c \cos\varphi,$$

$$\lambda / E_c = \cos\varphi = (-1/n) - 1.$$

Получили, что $\lambda$ определяется числом сигналов. Если используем два сигнала, то $\lambda/E_c = -1$, т. е. $\lambda = -E_c$. При $n = 3$ $\lambda = -E_c/2$. С увеличением $n$ коэффициент корреляции увеличивается, а поэтому ухудшаются условия передачи элементарных сигналов.

Связь между кодовым расстоянием и энергией сигнала позволяет установить условие целесообразности введения избыточности в код.

Пусть задан канал связи, который характеризуется шириной пропускания $F_к$, длительностью $T_к$ и превышением $H_к = \log_2 P_с/P_п$. Введем избыточность в код. Возможны два случая: 1) изменяется неизбыточный код, содержащий $m$ элементов. При введении избыточности добавляем $k$ контрольных элементов. В результате увеличивается время передачи; 2) время передачи не меняется, но уменьшается длительность элемента кода.

Определим оптимальное число элементов $k$. Пусть имеются две кодовые комбинации $y_{0j}$ и $y_{0i}$, тогда кодовое расстояние

$$d = \sqrt{\int_0^T (y_{0j} - y_{0i})^2 \, dt}.$$

В неизбыточном коде минимальное различие между кодовыми комбинациями имеется лишь в одном разряде, поэтому

$$d_{ij} = \sqrt{\int_0^{T/m} [y_0(t) - y_1(t)]^2 \, dt},$$

где $y_0(t)$ и $y_1(t)$ — символы кода. Очевидно, что

$$[y_0(t) - y_1(t)]^2 = \alpha P_с,$$

где $\alpha$ — постоянный коэффициент. Тогда $d_{ij} = \sqrt{\alpha P_с T/m}$. Для избыточного кода имеется различие в $d_{\min}$ символах. Отсюда

$$d_{ij_и} = \sqrt{d_{\min} \int_0^{T/n} [y_0(t) - y_1(t)]^2 \, dt} = \sqrt{d_{\min} \alpha P_с T/n}.$$

Условие целесообразного использования избыточного кода имеет вид $d_{ij_и} > d_{ij}$, тогда $d_{\min} \alpha P_с T/n > \alpha P_с T/m$, или $d_{\min} m > n$, или $d_{\min} m > m + к$, так как $n = m + k$. Следовательно, $k < m(d_{\min} - 1)$. Если данное условие соблюдается, то избыточный код дает эффект по помехоустойчивости, и чем больше информационная часть кода, тем большие значения $k$ могут иметь место для того же $d_{\min}$.

## § 2. Выделение детерминированного сигнала на фоне нормального шума

Пусть имеется канал связи, в котором действует помеха с амплитудой, подчиняющейся нормальному закону распределения, и передается код с двумя значениями сигнального признака $x_1(t)$, $x_2(t)$. Требуется определить вероятность правильного выявления $x_1(t)$ и $x_2(t)$. Пусть на рис. 63 изображена реализация помехи $n(t)$. Считаем, что значения амплитуд являются коррелированными на интервале $\Delta t$. Отложим на графике ряд таких интервалов корреляции. Значения $n_1$,

$n_2, \ldots, n_k, \ldots$ взаимно независимы, и поэтому можно заменить непрерывное распределение амплитуд дискретным и ввести функцию

$$f(n_1, n_2, \ldots, n_k, \ldots, n_n) = \left(1/\sigma_n \sqrt{2\pi}\right) \exp\left[-\sum_{k=1}^{n} n_k^2(t)/2\sigma_n^2\right],$$

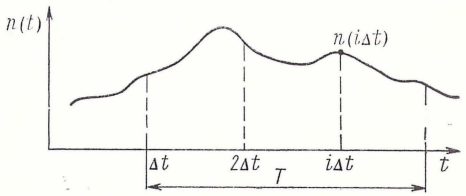


Рис. 63.

которая определяет вероятность того, что помеха примет одно из дискретных значений. Пусть на приемную сторону поступает сигнал $y(t)$, который складывается из переданного сигнала $x_1(t)$ и значения амплитуды помехи в момент передачи сигнала $n_1(t)$, т. е. $y(t) = x_1(t) + n_1(t)$. При передаче второго сигнала $y(t) = x_2(t) + n_2(t)$. Определим вероятности $p[x_1(t)/y(t)]$ и $p[x_2(t)/y(t)]$. Сравнив их, отдадим предпочтение тому из выявленных сигналов, для которого вероятность больше. Подставим в функцию $f$ значение $\Delta t$:

$$f(n_1, n_2, \ldots, n_k, \ldots, n_n) =$$
$$= \left(1/\sigma_n \sqrt{2\pi}\right) \exp\left\{-\sum_{k=1}^{n} \left[n_k^2(t)\,\Delta t/2\sigma_n^2 \Delta t\right]\right\}.$$

Наибольшее значение помеха имеет в момент независимого отсчета. Спектральная мощность помехи $G_\text{п} = \sigma_\text{п}^2 \Delta t$.

Рассмотрим $\Phi[n(t)] = \lim\limits_{\Delta t \to 0} f(n_1, n_2, \ldots, n_k, \ldots, n_n) =$

$$= \lim_{\Delta t \to 0} \left(1/\sigma_n \sqrt{2\pi}\right) \exp\left\{-\sum_{k=1}^{n} (1/2G_n)\, n_k^2(t)\,\Delta t\right\} =$$

$$= \left(1/\sigma_n \sqrt{2\pi}\right) \exp\left\{-(1/2G_\text{п}) \int_0^\tau n^2(t)\,dt\right\},$$

тогда $\Phi[n_1(t)] = \left(1/\sigma_n \sqrt{2\pi}\right) \exp\left\{-(1/2G_\text{п}) \int_0^\tau [y(t) - x_1(t)]^2 dt\right\}$,

$$\Phi[n_2(t)] = \left(1/\sigma_n \sqrt{2\pi}\right) \exp\left\{-(1/2G_\text{п}) \int_0^\tau [y(t) - x_2(t)]^2\,dt\right\}.$$

Найдем отношение

$$\alpha = \Phi[n_1(t)]/\Phi[n_2(t)] =$$
$$= \exp\left[-(1/2G_\text{п}) \int_0^\tau \{[y(t) - x_1(t)]^2 - [y(t) - x_2(t)]^2\}\,dt =\right.$$
$$= \exp\left[-(1/2G_\text{п}) \left\{\int_0^\tau [x_1(t)]^2\,dt - \int_0^\tau [x_2(t)]^2\,dt -\right.\right.$$
$$\left.\left. - \int_0^\tau 2y(t)\,[x_1(t) - x_2(t)]\,dt\right\}\right].$$

В этой формуле $\int_0^\tau [x_1(t)]^2\,dt$ — энергия первого сигнала, $\int_0^\tau [x_2(t)]^2\,dt$ — энергия второго сигнала.

В дальнейшем будем считать, что сигналы имеют одинаковые амплитуды, т. е. одинаковы по энергиям, тогда

$$\alpha = \exp\{(1/G_\text{п})\int_0^\tau y(t)\,[x_1(t) - x_2(t)]\,dt.$$

Обозначим $\theta = \int_0^\tau y(t)\,[x_1(t) - x_2(t)]$, тогда $\alpha = \exp(\theta/G_n)$. На приемной стороне вместо значения $\alpha$ будем определять значение $\theta$. Достаточно определить лишь знак $\theta$: если $\theta > 0$, то $\alpha > 1$, т. е. имеет место сигнал $x_1(t)$; если $\theta < 0$, то $\alpha < 1$ и имеет место сигнал $x_2(t)$. При отсутствии сигнала $y(t) = n(t)$, $\theta = 0$.

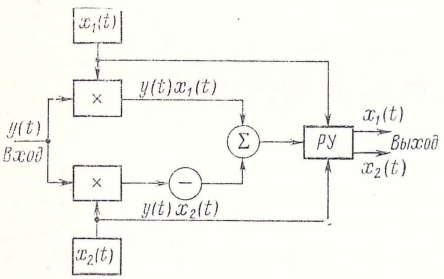


Рис. 64.

На рис. 64 представлена схема приема детерминированного сигнала. Значение сигнала $y(t)$, принятого из канала связи, перемножается со значениями сигналов $x_1(t)$ и $x_2(t)$, которые формируются генераторами на приемной стороне. С помощью сумматора осуществляется интегрирование на интервале $\tau$. В зависимости от знака результата в решающем устройстве (РУ) на выход отформировывается либо первый, либо второй сигнал.

Оценим вероятность ошибки при использовании данного способа. Пусть передавался сигнал $x_1(t)$, тогда $\theta$ должно быть положительно, но при расчете получилось $\theta < 0$, т. е. считаем, что поступил сигнал $x_2(t)$ с вероятностью $p(x_2/x_1)$. Если передавался сигнал $x_2(t)$, но получили $\theta > 0$, то поступил $x_1(t)$ с вероятностью $p(x_1/x_2)$.

Введем закон распределения значений $\theta$. Обозначим математическое ожидание величины $\theta$ через $\bar{\theta}$, дисперсию — через $\sigma_\theta^2$. Поскольку амплитуда помехи подчинена нормальному закону распределения, то используем нормальный закон для оценки $\theta$. Тогда

$$p(x_1/x_2) = (1/\sigma_\theta\sqrt{2\pi})\int_0^\infty \exp[-(\theta - \bar{\theta})^2/2\sigma_\theta^2]\,d\theta,$$

$$p(x_2/x_1) = (1/\sigma_\theta\sqrt{2\pi})\int_{-\infty}^0 \exp[-(\theta - \bar{\theta})^2/2\sigma_\theta^2]\,d\theta.$$

Если передается сигнал $x_2(t)$, то

$$\theta = \int_0^\tau y(t)\,[x_1(t) - x_2(t)]\,dt,\ y(t) = x_2(t) + n_2(t).$$

Математическое ожидание

$$\bar{\theta} = M\left\{\int_0^\tau y(t)\,[x_1(t) - x_2(t)]\,dt\right\} =$$

$$= M\left\{\int_0^\tau x_2(t)\,x_1(t)\,dt - \int_0^\tau [x_2(t)]^2\,dt + \right.$$

$$\left. + \int_0^\tau n_2(t)\,[x_1(t) - x_2(t)]\,dt\right\} = \lambda - E_c,$$

так как $\int_0^\tau x_2(t)\,x_1(t)\,dt = \lambda$, $\int_0^\tau [x_2(t)]^2\,dt = E_c$.

Дисперсия $\sigma_\theta^2 = M[\theta - \bar{\theta}]^2$,

$$\theta - \bar{\theta} = \int_0^\tau [x_2(t) + n_2(t)]\,[x_1(t) - x_2(t)]\,dt - \bar{\theta} =$$

$$= \int_0^\tau x_2(t)\,x_1(t)\,dt - \int_0^\tau [x_2(t)]^2\,dt + \int_0^\tau n_2(t)\,[x_1(t) -$$

$$- x_2(t)]\,dt - \bar{\theta} = \lambda - E_c + \int_0^\tau n_2(t)\,[x_1(t) - x_2(t)]\,dt -$$

$$- (\lambda - E_c) = \int_0^\tau n_2(t)\,[x_1(t) - x_2(t)]\,dt,$$

$$(\theta - \bar{\theta})^2 = \left\{\int_0^\tau n_2(t)\,[x_1(t) - x_2(t)]\,dt\right\}^2 =$$

$$= \int_0^\tau \int_0^\tau n_2(t)\,n_2(u)\,[x_1(t) - x_2(t)]\,[x_1(u) - x_2(u)]\,dt\,du,$$

где $u$ — некоторый аргумент, имеющий размерность времени и изменяющийся в интервале от 0 до $\tau$, тогда

$$\sigma_\theta^2 = M\int_0^\tau \int_0^\tau n_2(t)\,n_2(u)\,[x_1(t) - x_2(t)]\,[x_1(u) - x_2(u)]\,dt\,du,$$

$n_2(t)\,n_2(u) = G_n\delta(t-u)$, где $G_n$ — спектральная мощность помехи, $\delta(t-u)$ — импульсная функция первого порядка. Далее,

$$\sigma_\theta^2 = M\left\{G_n \int_0^\tau \int_0^\tau \delta(t-u)\,[x_1(u) - x_2(u)]\,du \times\right.$$

$$\left. \times\,[x_1(t) - x_2(t)]\,dt\right\}.$$

Так как $\int_0^\tau [x_1(u) - x_2(u)]\,\delta(t-u)\,du = x_1(t) - x_2(t)$, то

$$\sigma_\theta^2 = M\left\{G_n \int_0^\tau [x_1(t) - x_2(t)]^2\,dt\right\} =$$

$$= M\left\{G_n\left[\int_0^\tau [x_1(t)]^2\,dt - 2\int_0^\tau x_1(t)\,x_2(t)\,dt + \right.\right.$$

$$\left.\left. + \int_0^\tau [x_2(t)]^2\,dt\right\} = 2G_n[E_c - \lambda].\right.$$

Зная $\sigma_\theta$, определим

$$p(x_1/x_2) = \left(1/\sigma_\theta\sqrt{2\pi}\right)\int_0^\infty \exp\left[-(\theta - \bar{\theta})^2/2\sigma_\theta^2\right]\,d\theta.$$

Введем переменную $t = (\theta - \bar{\theta})/\sigma_\theta$, тогда

$$p(x_1/x_2) = \left(1/\sqrt{2\pi}\right)\int_{-\bar{\theta}/\sigma_\theta}^\infty \exp(-t^2/2)\,dt = \left(1/\sqrt{2\pi}\right)\left[1 -\right.$$

$$\left. - \int_0^{-\bar{\theta}/\sigma_\theta} \exp(-t^2/2)\,dt\right] = \left(1/\sqrt{2\pi}\right)[1 - \Phi(-\bar{\theta}/\sigma_\theta)].$$

Так как $\bar{\theta} = \lambda - E_c$, $\sigma_\theta^2 = 2G_п(E_c - \lambda)$, то

$$-\bar{\theta}/\sigma_\theta = (E_c - \lambda)/\sqrt{2G_п(E_c - \lambda)} = \sqrt{E_c(1 - \lambda/E_c)/2G_п}.$$

Окончательно получаем

$$p(x_1/x_2) = (1/\sqrt{2\pi})\left[1 - \Phi\left\{\sqrt{E_c(1 - \lambda/E_c)/2G_п}\right\}\right].$$

Из формулы следует, что для снижения вероятности трансформации необходимо выбирать по возможности большее отношение $E_c/G_п$, причем чем больше $E_c$, тем меньше вероятность ошибки. Целесообразно выбирать коэффициент корреляции $\lambda < 0$.

## § 3. Выделение случайного сигнала на фоне нормального шума

Этот процесс имеет место в системах телеизмерения, в автоматизированных системах управления при снятии непрерывной информации с датчиков и передаче этой информации на малые расстояния. Сигнал, отображающий непрерывную информацию, назовем случайным. Выделение его на фоне помех осуществляется с помощью построения оптимальных фильтров, обладающих некоторыми заданными параметрами.

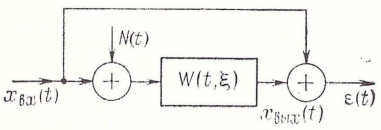

Рис. 65.

На рис. 65 изображена модель оптимального фильтра. Приняты следующие обозначения: $x_{вх}(t)$, $x_{вых}(t)$ — сигналы на входе и на выходе фильтра, $N(t)$ — помеха с заданным распределением амплитуд, $W(t, \xi)$ — импульсная переходная характеристика. Необходимо, чтобы величина $\varepsilon(t) = x_{вх}(t) - x_{вых}(t)$ была минимальной.

Определим оптимальную переходную характеристику $W(t, \xi)$. Сигнал

$$x_{вых}(t) = \int_0^t W(t, \xi)\, x_1(\xi)\, d\xi$$

есть реакция фильтра, возникающая в момент времени $\xi = 0$ на сигнал $x_{вх}(t)$. Тогда

$$\varepsilon^2(t) = x_{вх}^2(t) - 2x_{вх}(t)\int_0^t W(t, \xi)\, x_1(\xi)\, d\xi + \int_0^t \int_0^t W(t, \xi)\, W(t, \lambda)\, x_1(\xi)\, x_1(\lambda)\, d\xi d\lambda.$$

Определим среднеквадратичное значение ошибки фильтрации:

$$M[\varepsilon^2(t)] = \sigma^2(t) = M\left\{x_{вх}^2(t) - 2\int_0^t W(t, \xi)\, k_{1вх}(\xi, t)\, d\xi + \int_0^t \int_0^t W(t, \xi)\, W(t, \lambda)\, k_{11}(\xi, \lambda)\, d\xi d\lambda\right\},$$

где $k_{1вх}(\xi, t)$, $k_{11}(\xi,\lambda)$ — корреляционные функции. Можно подобрать $W(t, \xi)$ таким образом, чтобы получить ошибку $\sigma^2$ минимальной.

Пусть имеется некоторая оптимальная переходная характеристика $W_{опт}(t, \xi)$, удовлетворяющая минимуму $\sigma^2$. Очевидно, что

$$W(t, \xi) = W_{опт}(t, \xi) + \nu h(t, \xi),$$

где $\nu h(t, \xi)$ — отличие $W(t, \xi)$ от оптимальной. Тогда

$$\sigma^2(t) = M\{x^2_{вх}(t) - 2\int_0^t W_{опт}(t, \xi)\, k_{1вх}(\xi, t)\, d\xi +$$

$$+ \int_0^t \int_0^t W_{опт}(t, \xi)\, W_{опт}(t, \lambda)\, k_{11}(\xi, \lambda)\, d\xi d\lambda -$$

$$- 2\nu \int_0^t h(t, \xi)\, k_{1вх}(\xi, t)\, d\xi + \nu^2 \int_0^t \int_0^t h(t, \xi)\, h(t,\lambda)\, k_{11}(\xi, \lambda)\, d\xi\, d\lambda +$$

$$+ 2\nu \int_0^t \int_0^t W_{опт}(t, \lambda)\, h(t, \xi)\, k_{11}(\xi, \lambda)\, d\xi d\lambda\},$$

$$\int_0^t \int_0^t h(t, \xi)\, h(t, \lambda)\, M\{x_1(\xi)\, x_1(\lambda)\}\, d\xi d\lambda =$$

$$= M\left\{\int_0^t h(t, \xi)\, x_1(\xi)\, d\xi\right\}^2 = C^2.$$

Окончательно получаем

$$\sigma^2(t) = \sigma^2_{min} - 2\nu \int_0^t h(t, \xi)\, k_{1вх}(\xi, t)\, d\xi +$$

$$+ \nu^2 c^2 + 2\nu \int_0^t \int_0^t W_{опт}(t, \lambda)\, h(t, \xi)\, k_{11}(\xi, \lambda)\, d\xi d\lambda.$$

Возьмем производную $\partial\sigma^2(t)/\partial\nu$ и приравняем ее нулю:

$$\partial\sigma^2(t)/\partial\nu = -2\int_0^t h(t, \xi)\, k_{1вх}(\xi, t)\, d\xi +$$

$$+ 2\nu c^2 + 2\int_0^t \int_0^t W_{опт}(t, \lambda)\, h(t, \xi)\, k_{11}(\xi, \lambda)\, d\xi d\lambda.$$

Это выражение отображает реакцию системы с импульсной переходной функцией $h(t, \xi)$ на входной сигнал

$$k_{1вх}(\xi, t) - \int_0^t W_{опт}(t, \lambda)\, k_{11}(\xi, \lambda)\, d\lambda.$$

При $\nu = 0$

$$\int_0^t h(t, \xi)\left[k_{1вх}(\xi, t) - \int_0^t W_{опт}(t, \lambda)\, k_{11}(\xi, \lambda)\right] d\xi = 0.$$

Поскольку $h(t, \xi) \neq 0$, то

$$k_{1вх}(\xi, t) = \int_0^t W_{опт}(t, \lambda)\, k_{11}(\xi, \lambda)\, d\lambda.$$

Чтобы определить $W_{опт}(t, \lambda)$, необходимо знать корреляционную функцию входного сигнала и помехи, т. е. $k_{1вх}(\xi, t)$ и автокорреляционную функцию полезных входного и выходного сигналов $k_{11}(\xi, \lambda)$.

## § 4. Методы приема сигналов

Рассмотренные выше способы приема сигналов позволяют выявить любой элемент кода. Различают два основных способа приема сообщений: поэлементный и прием в целом.

*Поэлементный прием* осуществляется с использованием двух последовательно соединенных решающих схем. На рис. 66 представлена блок-схема приема. Пусть из линии связи поступает $n$-элементный код $y_1^{(1)}, y_2^{(2)}, \ldots, y_K^{(n)}$, где $y_1, \ldots, y_K$ — символы основания кода. Первая решающая схема (1РС) выявляет последовательно во времени элементы кода и запоминает их на время приема $n\tau$, вторая решающая схема (2РС) определяет номер принятого сообщения.

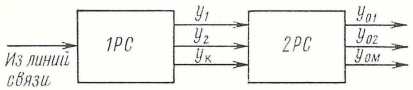


Рис. 66.

На рис. 67 изображен граф, характеризующий первую решающую схему. Пусть на входе имеет место последовательность символов $x_1, x_2, \ldots, x_K$, тогда 1РС выявляет последовательность $y_1, y_2, \ldots, y_K$. Правильность выявления определяется действием помех в канале связи, а также методом приема элементов кода. Вторая решающая схема работает в соответствии с графом, показанным на рис. 68. Она реализует обнаруживающие свойства кода.

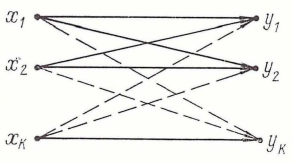


Рис. 67.

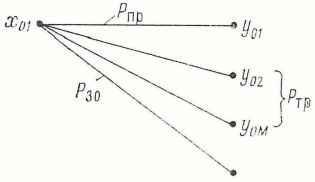


Рис. 68.

Достоинством поэлементного приема являются простота реализации и возможность практического использования различных типов кода. К недостаткам следует отнести неполное использование информации, содержащейся в символах кода. Рассмотрим пример реализации поэлементного приема для передачи четырех сообщений с использованием двоичного кода, обнаруживающего ошибку. Число кодовых переходов $d = r + s + 1$. Запишем кодовые комбинации 000, 011, 101, 110. Пусть требуется принять код третьего сообщения, представленный на рис. 69. Здесь $n$ — пуско-

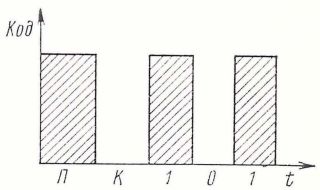


Рис. 69.

вой элемент кода, предназначенный для запуска схемы; $k$ — контрольный промежуток, позволяющий измерить длительность пускового элемента.

На рис. 70 показана схема поэлементного приема [30]. Код из линии связи через линейное устройство (ЛУ), согласовывающее канал связи с приемной частью системы, поступает на все схемы совпадения 1СС — 3СС и на элемент защиты пуска ЗП, с помощью которого после проверки пускового элемента по длительности осуществляется однократный запуск распределителя Р. Распределитель последовательно во времени с выходов 1 — 3 формирует импульсы, при этом параметры распределителя подобраны так, что импульсы синхронно располагаются посредине элементов выявляемого кода. Элементы кода выявляются схемами совпадения и приемными триггерами 1П — 3П. При кодовой комбинации 101 наблюдается два совпадения на схемах 1СС и 3СС, поэтому срабатывают приемные триггеры 1П, 3П и фиксируют единицу. Декомбинатор (ДК) по принятой комбинации элементов кода устанавливает, что пришло сообщение 101, и срабатывает исполнительный элемент (ЗИ).

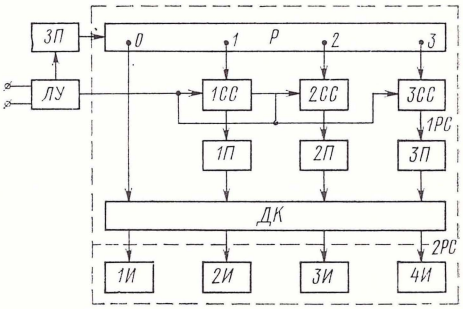

Рис. 70.

Метод *приема в целом* реализуется с использованием общей схемы выявления, в которой не представляется возможным выделить первую и вторую решающие схемы. При этом используется бóльшая информация в принятом сигнале относительно переданного сообщения, чем в поэлементном приеме, так как учитываются связи между отдельными символами в кодовой комбинации. Основная идея метода состоит в том, что решение принимается по всему коду в целом. Для принятия решения на приемной стороне оказывается необходимым генерировать всевозможные кодовые комбинации, которые входят в область допустимых значений, что в реальных случаях при реализации оптимальной схемы приема является весьма затруднительным. Поэтому прием в целом может быть реализован лишь тогда, когда на приемной стороне имеются достаточно мощные средства, представляющие собой специализированные вычислительные машины, которые позволяют осуществлять большой объем обработки информации.

Рассмотрим прием в целом сообщения, передаваемого двоичным кодом [22]. Предположим, что на вход приемного устройства поступил некоторый неизвестный нам сигнал, пред-

ставленный в виде кода $y'(t)$, состоящего из $n$ элементов длительностью $\tau$. Обозначим элементы кода $y^1(t)$, $y^2(t)$, ..., $y^n(t)$. Основная задача приемной схемы — определить номер выявляемого сообщения. Для этого сравним сигнал, принимаемый из линии связи, с любым возможным из $M$ допустимых сигналов. Выявление проведем в течение интервала времени $t = n\tau$, отбрасывая последовательно ненужные решения. Определим функцию корреляции между пришедшим сигналом $y'(t)$ и любым возможным сигналом из множества на основе суммирования отдельных функций корреляции, взятых по интервалам времени $\tau$, т. е. будем иметь $n$ слагаемых. Можно считать, что выявлено сообщение $x_{0i}$, если выполняется следующее неравенство:

$$\int_0^{\tau} y'(t)\, y_i^{(1)}(t)\, dt + \ldots + \int_{(j-1)\tau}^{j\tau} y'(t)\, y_i^{(j)}(t)\, dt + \ldots$$

$$\ldots + \int_{(n-1)\tau}^{n\tau} y'(t)\, y_i^{(n)}(t)\, dt > \int_0^{\tau} y'(t)\, y_q^{(1)}(t)\, dt + \ldots$$

$$\ldots + \int_{(j-1)\tau}^{j\tau} y'(t)\, y_q^{(j)}(t)\, dt + \ldots + \int_{(n-1)\tau}^{n\tau} y'(t)\, y_q^{(n)}(t)\, dt,$$

которое должно выполняться при всех значениях $g \neq i$. При этом под $y'(t)$ понимается сигнал, принятый из линии связи. Полученное неравенство можно переписать в виде

$$\sum_{j=1}^{n} \int_{(j-1)\tau}^{j\tau} y'(t)\, y_i^{(j)}(t)\, dt > \sum_{j=1}^{n} \int_{(j-1)\tau}^{j\tau} y'(t)\, y_q^{(j)}(t)\, dt.$$

Нетрудно показать, что такое же неравенство может быть записано и для обратных значений символов в элементах кода, т. е.

$$\sum_{j=1}^{n} \int_{(j-1)\tau}^{j\tau} y'(t)\, \overline{y_i^{(j)}(t)}\, dt > \sum_{j=1}^{n} \int_{(j-1)\tau}^{j\tau} y'(t)\, \overline{y_q^{(j)}(t)}\, dt,$$

где $\overline{y_i^{(j)}(t)}$ — элемент, инверсный относительно $y_i^{(j)}(t)$. В приведенных выше неравенствах $i$ — номер сообщения, $j$ — порядковый номер элемента в коде сообщения. Так как сообщения передаются двоичными кодами, то инверсный элемент кода является однозначно определяемым.

Для определения алгоритма выявления любого сообщения вычтем из исходного неравенства неравенство, записанное для инверсных элементов:

$$\sum_{j=1}^{n} \int_{(j-1)\tau}^{j\tau} y'(t) \left[ y_i^{(j)}(t) - \overline{y_i^{(j)}(t)} \right] dt >$$

$$> \sum_{j=1}^{n} \int_{(j-1)\tau}^{j\tau} y'(t) \left[ y_q^{(j)}(t) - \overline{y_q^{(j)}(t)} \right] dt.$$

Введем обозначения $y_i^{(j)}(t) - \overline{y_i^{(j)}(t)} = y_{\Delta}(t)\, \beta_{ij}$, где $\beta_{ij}$ — коэффициент, определяющий знак приращения; $y_{\Delta}(t)$ — приращение с амплитудой, равной амплитуде единичного импульса.

Записанное выше неравенство соответствует $i$-му сообщению. Аналогично для $q$-го сообщения можно записать:

$$y_q^{(j)}(t) - \overline{y_q^{(j)}(t)} = y_\Delta(t)\,\beta_{dj}.$$

Тогда, используя полученные выше соотношения, находим

$$\sum_{j=1}^{n} \beta_{ij} \int_{(j-1)\tau}^{j\tau} y'(t) y_\Delta(t)\, dt > \sum_{j=1}^{n} \beta_{qj} \int_{(j-1)\tau}^{j\tau} y'(t)\, y_\Delta(t)\, dt,$$

$$\sum_{j=1}^{n} \beta_{ij} c_j > \sum_{j=1}^{n} \beta_{qj} c_j.$$

Полученное неравенство позволяет формальным образом оценить номер выявляемого сообщения и, следовательно, построить схему приема в целом.

Схема приема в целом представлена на рис. 71. Значение неизвестного сигнала вида $y'(t)$ из линии связи перемножается на приемной стороне со значением сигнала $y_\Delta(t)$, который выдается формирователем Ф и воспроизводится с его выхода с тактом $\tau$. С помощью включенных в схему умножителя и интегратора осуществляется выявление коэффициентов $c_j$. В схеме предусмотрено наличие памяти (П), где осуществляется хранение набора всевозможных коэффициентов $\beta_{ij}$, $\beta_{qj}$, характеризующих коды всех возможных сообщений. Коэффициенты, полученные на выходе интегратора, перемножаются с соответствующими коэффициентами, находящимися в памяти, с помощью умножителей, и поступают далее на устройство сравнения, осуществляющее реализацию алгоритма выявления, т. е. устанавливающее номер сообщения, для которого удовлетворяется основное неравенство. С выхода устройства сравнения по одной из выходных шин выдается сигнал выявления данного сообщения.

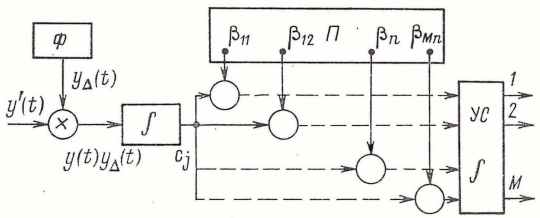


Рис. 71.

Особенностью приема в целом является то, что используется практически вся информация в принятом сигнале о переданном сообщении, а поэтому, как следствие, мы получаем наилучшие показатели передачи информации по помехоустойчивости. Недостатками метода можно считать сложность

реализации, необходимость наличия вычислительных и запоминающих устройств на приемной стороне. Это является справедливым, если не учитывать наличия обратного канала. Л. М. Финком было показано [22], что если сравнивать прием в целом с поэлементным, полагая, что при поэлементном приеме используется обнаружение ошибок и реализуется система с запросом по идеальному каналу обратной связи, то прием в целом может уступать по помехоустойчивости поэлементному, поэтому в настоящее время при передаче информации в АСУ весьма широкое распространение нашел поэлементный прием сигналов с построением кодов, рассчитанных на обнаружение ошибок и передачу сигнала обратной связи в виде кода запроса или квитанции.

В связи со значительной сложностью реализации приема в целом в ряде случаев находят применение некоторые варианты приема, занимающие по помехоустойчивости промежуточные положения между поэлементным приемом и приемом в целом. При этом используются также две решающие схемы, т. е. сначала преобразуется принятая последовательность в последовательность элементов кода, а далее эта последовательность элементов кода декодируется. Особенность построения таких схем состоит в использовании апостериорной вероятности выявления символов, которая учитывается в процессе декодирования.

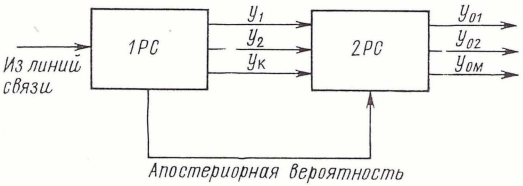


Рис. 72.

Примером указанного метода приема является прием по наиболее надежным символам, структурная схема которого изображена на рис. 72. Данный вариант предусматривает наличие как первой, так и второй решающих схем. Первая решающая схема осуществляет определение апостериорных вероятностей выявления символов и принимает предварительное решение о том, какой символ передавался при условии, что он имеет наибольшую апостериорную вероятность. С выхода первой решающей схемы получаем набор символов кода, выбранных на основе этого правила. Вторая решающая схема на основе выявленных символов осуществляет декодирование, т. е. определяет номер переданного сообщения. Отличие от поэлементного приема при этом состоит в том, что на вторую решающую схему подаются дополнительно апостериорные вероятности выявленных символов, а поэтому при выявлении кодовой комбинации можно ориентироваться лишь на те символы, кото-

рые имеют наибольшие апостериорные вероятности и являются наиболее надежными. Если между кодовыми комбинациями имеется кодовое расстояние $d$, то число наиболее надежных символов для отличия одной комбинации от другой может не превышать $n-d+1$.

Следует отметить, что в данном способе при выявлении сообщения используются информация, учитывающая корреляционную связь между символами, и апостериорные вероятности. Однако апостериорные вероятности касаются только наиболее надежных символов, а по менее надежным информация полностью теряется. В этом данный метод уступает методу приема в целом, а поэтому его характеристики по помехоустойчивости оказываются хуже. Можно показать, что при отдельных моделях помех и кодах с малой избыточностью прием в целом и прием по наиболее надежным символам могут быть эквивалентны в смысле верности передаваемой информации.

Несколько худшим по верности является метод Вагнера Структура построения его приемной схемы аналогична предыдущей и включает в себя первую и вторую решающие схемы. Метод Вагнера используется лишь для двоичных кодов с четным кодовым расстоянием, поэтому при поэлементном декодировании можно с помощью проверок на четность определить число ошибок. Если число ошибок составляет $d/2$, то информация об апостериорных вероятностях используется и наименее достоверный символ изменяется на противоположный. Если ошибка при этом исправляется, то число ошибок уменьшается до $(d/2)-1$. Оставшиеся ошибки исправляются путем линейных проверок на четность. Для кодов $d=2$ метод Вагнера и метод приема по наиболее надежным символам практически не отличаются друг от друга и обеспечивают ту же верность передачи, что и метод приема в целом. Следует отметить, что метод приема по наиболее надежным символам, предложенный Л. Ф. Бородиным [3], и метод Вагнера по схемной реализации являются более приемлемыми и реализуются проще, чем прием в целом, и тем самым могут найти практическое применение при передаче информации в АСУ.

## § 5. Выделение шумоподобных сигналов на фоне помех

Наилучшие условия согласования с помехой дает сигнал, который по своим статистическим свойствам не отличается от помехи. Если помеха представляет «белый» шум, то для передачи информации при условии $P_с/P_п \ll 1$ используются периодические сигналы в виде псевдослучайных последовательностей. Псевдослучайная последовательность — это набор импульсов разной полярности, характеризуемый числом

импульсов $n$, амплитудой импульса $a$ и функцией корреляции (см. [18]):

$$\lambda_j = \int_0^T x(t)\, x(t-\tau_j)\, dt,$$

где $T$ — время существования сигнала, $x(t)$ — мгновенное значение сигнала в момент времени $t$, $\tau_j$ — интервал корреляции, который обычно принимается кратным длительности элемента кода (импульса последовательности).

При $\tau = 0$ можно рассчитать автокорреляционную функцию $\lambda_0 = \int_0^T [x(t)]^2\, dt$. При $\tau_j = j\tau$ определяются корреляционные функции $\lambda_j\ (1 \leqslant j \leqslant n)$. Формирование псевдослучайной последовательности можно осуществить с использованием устройства, структурная схема которого представлена на рис. 73. Формирователь состоит из триггеров $T_1$, $T_2$, $T_3$, сумматора и генератора тактовых импульсов (ГТИ), обеспечивающего с тактом $\tau$ сдвиг символов, записанных в триггерах. Функционирование схемы осуществляется в соответствии с уравнением $x_0 = x_2 \oplus x_3$.

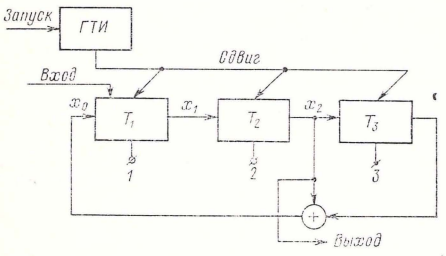

Рис. 73.

Введем оператор сдвига $x$. Будем считать, что сигнал $x_{i+1}$ есть сдвинутый на один шаг сигнал $x_i$. Операция сдвига условно записывается следующим образом: $x_{i+1} = xx_i$, $x_1 = xx_0$, $x_2 = xx_1 = x^2x_0$, $x_3 = xx_2 = x^3x_0$, и основное уравнение функционирования формирователя имеет вид

$$x_0 = x^3x_0 + x^2x_0.$$

Отсюда $x^3x_0 \oplus x^2x_0 \oplus x_0 = 0$, $x_0(x^3 \oplus x^2 \oplus 1) = 0$, $x^3 \oplus x^2 + 1 = p(x)$; $p(x)$ называется *порождающим многочленом*. Формирователь осуществляет умножение некоторого исходного сигнала $x_0$ на многочлен $p(x)$, определенный соединением цепей формирователя. Рассмотрим, как работает схема, если $x_0 = 1$.

Пусть все триггеры были сначала в нулевом положении. В некоторый момент времени подается импульс на вход первого триггера. Рассмотрим сигналы $x_1$, $x_2$, $x_3$, $x_0$ как последовательности во времени: 1011100 1011100 ..., 0101110 0101110 ..., 0010111 0010111 ..., 0111001 0111001 ...

В основе сигналов лежит одна и та же последовательность, например $x_2$ вида 0101110. Остальные получаются сдвигом относительно ее на определенное число тактов. Полученная последовательность $x_2$ является псевдослучайной и имеет дли-

ну $n = 2^m - 1$, где $m$ — число триггеров. Она позволяет обеспечить передачу сигнала при $P_с \ll P_п$.

Рассмотрим принцип выявления псевдослучайных последовательностей на фоне помех. Пусть последовательность $x_2$, имеющая вид 0101110, используется для передачи одного сообщения. Обеспечим многократную передачу $x_2$ и оценим на приемной стороне корреляционную функцию. Для этого установим связь между $x_2$ и сигналами такой же структуры, сдвинутыми относительно $x_2$ на число тактов от 0 до $n$. Тогда получаем набор значений функции $\lambda_j$. Нетрудно показать, что

$$\lambda_0 = 7a^2\tau = na^2\tau = a^2T = E_с,\ \lambda_j = -a^2\tau\ (1 \leqslant j \leqslant n),$$

а график корреляционной функции представлен на рис. 74. Для выявления передаваемого сообщения на приемной стороне устанавливаем измеритель корреляционной функции. Превышение функции над некоторым порогом означает прием сообщения. При наличии помех величина пика автокорреляционной функции уменьшается, поэтому порог должен выбираться с учетом действующих помех. Недостатком этого способа является то, что для выдачи каждого сообщения требуется свой формирователь, поэтому структура системы с ростом числа передаваемых сообщений резко усложняется.

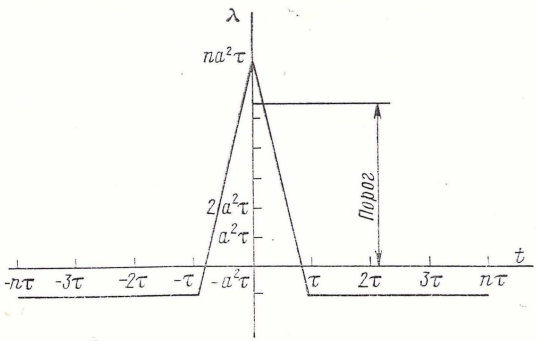


Рис. 74.

При передаче информации в АСУ псевдослучайные последовательности могут использоваться для построения избыточных циклических кодов, а также для передачи служебной информации. Выявим наиболее эффективный путь их использования. Для этого оценим пропускную способность и скорость передачи информации в канале связи с шумоподобными сигналами.

Пусть формируется псевдослучайная последовательность длиной $n$, отображающая $M$ сообщений, причем $M = n$.

Пропускная способность канала связи в условиях действия нормального шума

$$C = F\log_2(1 + P_c/P_п).$$

Длительность сигнала $T = n\tau$. Введем понятие длительности передачи одной двоичной единицы информации $\tau_0$, тогда $T = \tau_0 I$, где $I$ — количество информации, содержащееся в одной псевдослучайной последовательности. Если передаваемые сообщения равновероятны, то $I = \log_2 M$ и для двоичного кода, отображающего сообщения, $T = \tau_0 \log_2 M = \tau_0 \log_2 n$, $\tau = \tau_0 \times (\log_2 n)/n$. Известно, что ширина полосы пропускания $F = 1/2\tau$, поэтому $F = n/2\tau_0 \log_2 n$. Мощность помехи $P_п = G_п F$, где $G_п$ — мощность помехи, приходящаяся на единицу ширины полосы пропускания канала связи. Пропускная способность

$$C = (n/2\tau_0 \log_2 n)\log_2[1 + (2P_c\tau_0 \log_2 n/G_п n)].$$

Введем коэффициент превышения по энергии сигнала над помехой: $\alpha = P_c\tau_0/G_п$. Тогда $C\tau_0 = \log_2[1 + (2\alpha(\log_2 n)/n)]^{n/2\log_2 n}$. Из формулы следует, что псевдослучайную последовательность целесообразно применять при большом числе передаваемых сообщений. Если использовать ее для передачи одного сообщения, то получим очень малую пропускную способность, и для обеспечения достоверной передачи необходимо время, намного превышающее $T$. Учитывая это, для выдачи избыточного кода используем формирователь, представленный на рис. 73. Тогда число используемых триггеров соответствует числу информационных позиций в коде $m$ и возможно построение избыточного кода для числа сообщений $M = 2^m$ при длине кода $n = 2^m - 1$. В частном случае при $m = 3$, $n = 7$ получаем код с кодовым расстоянием $d = 4$. Формирование кодовых комбинаций осуществляется при подаче на входы 1, 2, 3 символов информационной части кода. Легко убедиться в том, что формируются следующие комбинации: 0010111, 0101110, 1001011, 1011100, 1100101, 1110010.

В общем случае кодовое расстояние для последовательности длиной $n = 2^m - 1$ составляет $d = (n + 1)/2$.

*Глава VII*

# ИНФОРМАЦИОННЫЙ ПОДХОД К ОЦЕНКЕ КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ КОМПЛЕКСА ТЕХНИЧЕСКИХ СРЕДСТВ АСУ

## § 1. Основные критерии качества функционирования комплекса технических средств АСУ

Разработка комплекса технических средств АСУ осуществляется на этапе внутреннего проектирования, когда определены цели, создана модель объекта, выбраны основные критерии эффективности и оптимальная стратегия управления. Целью внутреннего проектирования является построение обеспечивающих подсистем АСУ, основную техническую базу которых составляет КТС. Разработка комплекса технических средств может осуществляться только при условии, если выбраны основные критерии эффективности функционирования АСУ в целом и в этих критериях в качестве аргументов выступают критерии эффективности функционирования КТС. Поскольку основными задачами функционирования комплекса технических средств являются сбор, передача, обработка, отображение и выдача информации, т. е. различные фазы ее преобразования, то для каждой подсистемы КТС могут быть выбраны общие критерии оценки качества функционирования, объединяемые в обобщенный критерий качества функционирования всего комплекса.

Обобщенным критерием выбора оптимального варианта построения КТС является минимизация затрат. Обычно оценивают приведенные затраты на создание и эксплуатацию КТС: $C=A+kB$, где $A$ — затраты, распределенные во времени, $B$ — единовременные капитальные затраты, $k$ — коэффициент приведения единовременных затрат к расходам, распределенным во времени [1]. Коэффициент приведения $k$ зависит от типа отрасли промышленности, где внедряются АСУ. Обычно $0{,}33 \geqslant k \geqslant 0{,}2$. Рассмотрим составляющие затрат. Затраты на эксплуатацию складываются из удельных эксплуатационных затрат на единицу оборудования $i$-го типа $A_i$ и затрат на содержание административно-управленческого персонала, который необходим для эксплуатации АСУ, $Q_{\text{ау}}$. Отсюда $A=\Sigma_{i=1}^{m} A_i n_i + Q_{\text{ау}}$, где $n_i$ —

число устройств $i$-го типа, входящих в КТС, $m$ — число видов устройств. Капитальные затраты $B$ складываются из капитальных затрат по $i$-м техническим средствам $B_i$, т. е. $B=\Sigma_{i=1}^{n} B_i$, где $n$ — число видов технических средств. Величина $B_i$, определяемая для каждого технического средства, зависит от производственных затрат, стоимости стандартных и создаваемых оригинальных технических средств, стоимости строительства и реконструкции зданий, а также вспомогательного оборудования и монтажа технических средств.

Таким образом, обобщенный критерий приведенных затрат является основным критерием, который позволяет сравнить разные варианты и выбрать из них вариант, минимизирующий величину $C$. Однако этот критерий является правомочным, если для каждого варианта построения КТС выполняются основные ограничения по критериям, характеризующим качество преобразования информации в каждой подсистеме. Критериями качества преобразования информации можно считать верность и время преобразования информации. Они существенно зависят от назначения АСУ, выбранного варианта КТС и т. д. Обычно эти показатели должны находиться в заданных пределах. В простейшей методике расчета не учитывают особенностей объединения отдельных технических средств, наличия внутренней избыточности в структуре и избыточности в преобразуемой информации. Зачастую принимают, что каждая подсистема КТС содержит набор одинаковых однотипных технических средств, работающих как последовательное соединение элементов, и все подсистемы между собой соединены последовательно [17]. Тогда вероятность ошибки при преобразовании информации

$$P_{\text{о}} = 1 - \Pi_{i=1}^{n} (1 - p_{i\,\text{ПРПИ}}) \Pi_{j=1}^{m} (1 - p_{j\,\text{ПСПИ}}) \times$$
$$\times \Pi_{\xi=1}^{k} (1 - p_{\xi\,\text{ПХОИ}}) \Pi_{q=1}^{r} (1 - p_{q\,\text{ПВВИ}}),$$

где предполагается наличие $n$ технических средств ПРПИ, $m$ технических средств для ПСПИ, $k$ устройств ПХОИ, $r$ устройств ПВВИ; величины $p$ определяют вероятности возникновения ошибок в соответствующих технических средствах.

Следует отметить, что такой подход дает грубую оценку верности преобразования информации в подсистемах КТС, а поэтому необходим более глубокий анализ как причин, вызывающих искажения информации в АСУ, так и их последствий, т. е. ошибок, возникающих в преобразуемой информации. Соответственно время прохождения экономической информации в АСУ $T=T_{\text{КТС}}+T_{\text{р}}$, где $T_{\text{КТС}}$ — время преобразования информации всем комплексом технических средств, $T_{\text{р}}$ — время реакции управляющего персонала на воздействие со стороны АСУП.

Обычно эти составляющие раскрываются с учетом всех возможных подсистем путем суммирования времен, затрачиваемых на преобразование информации каждым техническим устройством подсистемы КТС.

Остановимся более подробно на подсистеме сбора и передачи информации в АСУ, так как в ней наиболее полно удается использовать основные результаты теории информации для обеспечения безошибочной передачи сообщений. Основным критерием, характеризующим качество передачи информации в данной подсистеме, может служить вероятностный критерий, оценивающий правильность выявления сообщения, т. е. вероятность $P_{\text{пр}}$. Эта вероятность совместно с вероятностью трансформации (преобразования) сообщения $P_{\text{тр}}$ и вероятностью защитного отказа $P_{\text{зо}}$ определяет полную группу событий, т. е. $P_{\text{пр}}+P_{\text{тр}}+P_{\text{зо}}=1$. Обеспечить заданные пределы этих показателей, т. е. $P_{\text{пр}}\geqslant[P_{\text{пр}}]_{\text{доп}}$, $P_{\text{тр}}\leqslant[P_{\text{тр}}]_{\text{доп}}$, можно путем введения избыточности в передаваемую информацию. Независимо от способов введения избыточности, что будет рассмотрено далее, наличие избыточности вызывает увеличение времени передачи сообщения, а поэтому существенным критерием, определяющим быстродействие системы, является скорость передачи информации. Выше было показано, что независимо от типа сообщения скорость передачи информации можно определить как количество информации, передаваемое в одном символе кода: $I=H(Y)-H(Y/X)$. Для равновероятных сообщений при наличии $m$ информационных и $k$ контрольных символов в коде нетрудно получить, что скорость передачи информации $I=(\log_2 M)/n=m(\log_2 K)/n$ определяется в основном долей информационных позиций в коде и основанием кода. Эта формула справедлива в частном случае, когда не используются дополнительные методы повышения помехоустойчивости передачи за счет обратного канала.

Независимо от способа введения избыточности полезно использовать критерий эффективности для подсистемы сбора и передачи информации в виде скорости передачи неискаженной информации. Тогда $I_{\text{и}}=IP_{\text{пр}}$ [3]. Очевидно, что использование данного критерия непосредственно невозможно, поскольку необходимо повышать скорость при заданных ограничениях по вероятностным показателям качества, а поэтому данный критерий может быть использован при дополнительном ограничивающем условии $P_{\text{пр}}\geqslant[P_{\text{пр}}]_{\text{доп}}$.

Следует отметить, что использование критерия скорости передачи неискаженной информации позволяет достаточно просто сравнивать различные методы передачи информации, которые отличаются длиной блока сообщений, объемом информации, уровнем и типом вводимой избыточности. Если по данному критерию при соответствующих ограничениях получен ряд вариантов построения подсистемы сбора и передачи информации КТС,

то возникает задача выбора оптимального варианта, и здесь может быть использован критерий приведенных затрат, исходя из которого целесообразно выбрать вариант, имеющий наименьшую стоимость.

Частным критерием, который также непосредственно связан со стоимостью аппаратуры, может служить критерий объема аппаратуры системы. В этом случае осуществляется сравнение этого объема с тем минимально необходимым объемом аппаратуры, который позволяет получить заданное качество функционирования. Трудность использования понятия «объем аппаратуры» заключается в необходимости установления некоторой эталонной системы, для которой определяется минимально необходимый объем. Обычно минимально необходимый объем может быть установлен, если известны некоторые оптимальные способы введения избыточности и выполнено сравнение данного метода введения избыточности с известным оптимальным способом.

Указанные выше критерии характеризуют лишь процесс передачи информации в системе и могут быть использованы в предположении идеальной надежности аппаратуры и отсутствия отказов по причине перегрузки буферных устройств хранения информации. С учетом ненадежности аппаратуры и отказов в обслуживании пересекающихся информационных потоков может быть введен обобщенный вероятностный критерий, оценивающий качество преобразования информации, а именно вероятность потерь. Тогда для определенного участка передачи и предварительной обработки информации в АСУ можно записать, что общая вероятность потерь

$$P_{\text{пот}} = P_{\text{пот.н}} + (1 - P_{\text{пот.н}})\, P_{\text{пот.о}},$$

где $P_{\text{пот.н}}$ — вероятность потерь, вызванных действием помех в линиях связи и отказов в аппаратуре подсистемы, $P_{\text{пот.о}}$ — вероятность потерь по причине отказа в обслуживании информационного потока.

Численный расчет данных составляющих зависит либо от принятой модели помех в канале связи, либо от соответствующей ей модели ошибок в выявляемой информации, а также от выбранного закона обслуживания информационных потоков, от метода введения и типа избыточности. В соответствии с этим рассмотрим более подробно верность преобразования информации в ПСПИ с учетом возможных мешающих воздействий.

## § 2. Информационный подход к оценке верности преобразования информации

Основной задачей подсистемы сбора и передачи информации в АСУ является передача информации с заданными быстродействием и верностью. На вход подсистемы поступает поток сооб-

щений, который далее может быть либо принят для передачи, либо не принят в связи с занятостью запоминающих или входных устройств данной подсистемы. В соответствии с этим поток сообщений, принятый для передачи, преобразуется в поток символов кода. При этом преобразовании также могут возникать определенные потери информации, вызванные ненадежностью формирующей аппаратуры, т. е. ненадежностью в основном кодирующих устройств и средств выдачи информации в канал связи. Поток символов, поступивший из канала связи на приемную часть ПСПИ, может быть принят на выявление и не выявлен по причине неисправности аппаратуры или не принят на выявление по причине ее занятости выявлением других информационных потоков. Однако даже если поток был принят на выявление, под действием помех в канале связи могут возникать такие ошибки, которые делают невозможным достоверное выявление информации. Последнее имеет место, если в данном цикле передачи введенной в информацию избыточности оказалось недостаточно для исправления ошибок, возникших под действием помех в канале связи. Из потока сообщений, поступающих на вход подсистемы, формируется некоторый поток потерянных сообщений (рис. 75). Независимо от места возникновения потерь информации основными причинами потерь в данной подсистеме являются неисправность аппаратуры, помехи в каналах связи и перегрузка обслуживающих (декодирующих) или запоминающих устройств. Рассмотрим эти причины.

Наличие неисправности в аппаратуре приводит к ее неработоспособности. Под надежностью любой информационной системы понимается свойство системы выполнять свои функции, «сохраняя во времени значения установленных эксплуатационных показателей в заданных пределах...»*. Так, при формировании кода на передающей стороне ПСПИ могут возникать искажения, вызванные ненадежностью аппаратуры. Необходимо учитывать возможность возникновения сбоев и отказов. Под сбоем обычно понимается «самоустраняющий отказ, приводящий к кратковременному нарушению работоспособности». Под отказом следует понимать нарушение работоспособности аппаратуры. Таким образом, отказы снимают функциональную зависимость между входом и выходом отдельных устройств в подсистемах, и для анализа их последствий необходимо иметь некоторую модель отказов. Обычно считают, что отказы являются случайными, внезапными и взаимно независимыми событиями.

Другим случайным воздействием, которое является весьма существенным в ПСПИ, является помеха. Под помехой обычно понимают возмущение, которое возникает в системе передачи

---

* См. ГОСТ 13377 — 75.

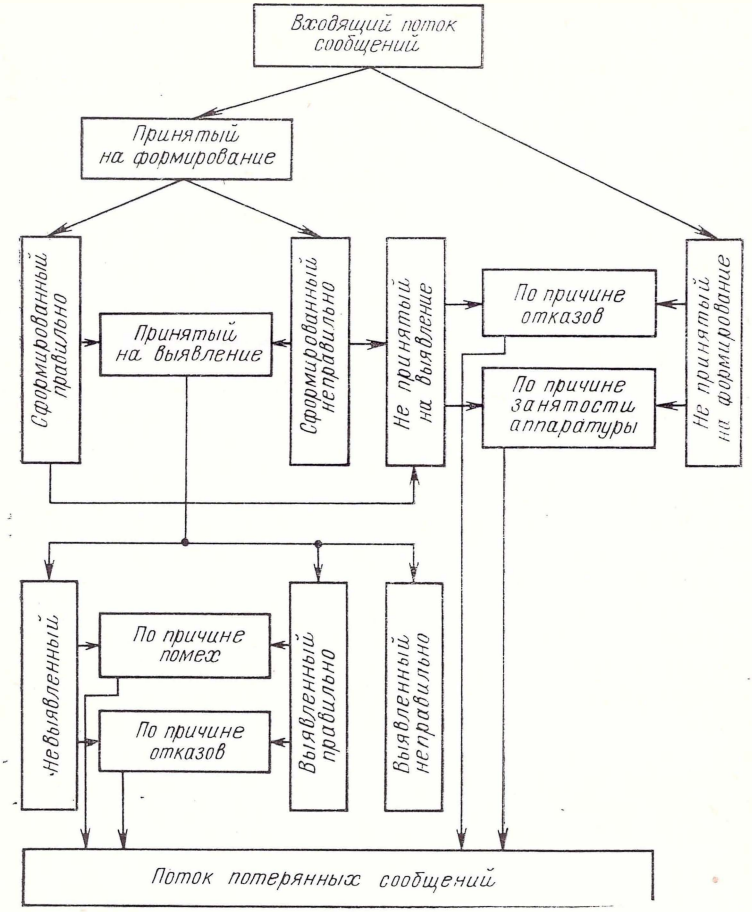


Рис. 75.

информации и препятствует правильному приему сигнала. Различают два основных вида помех: случайные и регулярные. Для АСУ регулярные помехи обычно нехарактерны, поэтому все расчеты, связанные с верностью передачи информации в условиях помехи для АСУ, следует проводить применительно к моделям случайных помех. При этом нередко удается отвлечься от понятия помехи, поскольку помеха связана с физическими процессами в канале связи, и исходить из последствий воздействия помехи на передаваемый код. Если помеха представляет собой некоторый исходный процесс, то ошибка есть следствие воздействия помехи на передаваемый код, и поэтому за исходные данные можно принять параметры некоторой модели ошибок в канале передачи информации. В качестве такой модели

на нижнем уровне передачи информации при наличии специально выделенных проводных линий связи может быть использована модель независимых ошибок. На верхних уровнях передачи, в частности для случая высокоскоростной передачи информации с использованием модемов, следует использовать модели, которые отображают группирование ошибок в отдельных временных интервалах, поскольку, как показывают расчеты, в данном случае использование модели независимых ошибок приводит к весьма грубым приближенным результатам.

При выявлении кода на приемной стороне основным случайным воздействием является неисправность аппаратуры, последствием которой могут быть сбои и отказы. К ним можно применить те же количественные характеристики оценки, что и на передающей стороне.

Под помехоустойчивостью будем понимать способность системы обеспечивать передачу информации с заданной верностью при воздействии помех в каналах связи. Основными критериями надежности и помехоустойчивости можно считать критерии, получаемые из статистического материала при большом числе испытаний элементов аппаратуры системы в целом, а также из анализа передачи и первичной обработки информации. Основным математическим аппаратом по оценке этих критериев является теория вероятности. В теории надежности установившимися критериями к настоящему времени являются вероятность безотказной работы, среднее время безотказной работы, интенсивность отказов, вероятность ошибки, вероятность потери информации и т. д.

Следует отметить, что применительно к информационным системам такие критерии, как вероятность безотказной работы, вероятность отказа, зачастую являются малопригодными. Более эффективными оказываются вероятность правильного формирования сигнала, вероятность трансформации, вероятность защитного отказа при формировании, если анализируется передающая часть системы. Соответственно для приемной части могут быть использованы такие вероятностные показатели, как вероятности правильного и неправильного выявления, вероятность защитного отказа при выявлении.

Положительные результаты могут быть получены, если применить информационный подход к оценке верности преобразования информации в подсистеме сбора и передачи. Прежде всего через энтропию можно оценить количество информации, которая передается и обрабатывается в данной подсистеме. Можно в единые критерии объединить оценку помехоустойчивости передачи информации и надежности функционирования технических средств, так как проблема обеспечения заданной надежности системы может быть решена теми же методами, что и задачи помехоустойчивости передачи информации.

Проблема надежности отличается от проблемы помехоустойчивости тем, что в случае отказа повторение одной и той же операции во времени не позволяет обнаружить и исправить ошибку. Вместе с тем в системах с последовательными кодами одиночный отказ элемента может привести к неодиночной ошибке в выходном сигнале. Однако при соответствующем проектировании информационных систем основные методы обеспечения помехоустойчивой передачи информации могут быть перенесены на методы конструирования надежных технических средств. В частности, с небольшими изменениями можно использовать большинство результатов теории кодирования. При этом необходимо решить ряд задач по оптимальному размещению и допустимой сложности декодирующего устройства, объему аппаратуры, который охватывается избыточностью.

Проблема помехоустойчивости в определенной степени является противоречивой по отношению к проблеме надежности. Если для увеличения помехоустойчивости необходимо увеличивать избыточность передаваемой информации, то это приводит в свою очередь к усложнению структуры системы, к увеличению ее объема, и если вводимая избыточность не рассчитывалась на исправление ошибок, возникших по причине неисправности аппаратуры, то наблюдается снижение надежности системы. Только оценивая помехоустойчивость и надежность единым критерием, можно определить общую эффективность построения данного варианта ПСПИ [6]. Если учесть, что основной задачей для ПСПИ является правильный прием переданного сообщения за допустимый промежуток времени, то можно воспользоваться такими едиными вероятностными критериями, как вероятность правильного приема $P_{пр}$, вероятность трансформации сообщения $P_{тр}$. Свойство системы в отношении помехоустойчивости и надежности можно связать с количеством информации, передаваемой через систему.

Представим обобщенно мешающие воздействия в виде некоторых условных канала с шумом отказов аппаратуры и канала с шумом помех в линии связи (рис. 76). Это соответствует выделению идеально функционирующей системы, в которой от-

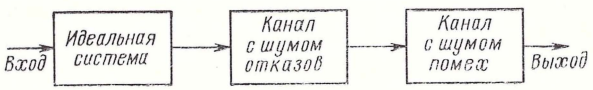


Рис. 76.

сутствуют потери информации, и последовательно связанного с ней канала с шумом, куда отнесены все возможные потери информации, возникающие под влиянием случайных воздействий в системе [4]. В этом случае полезно объединить каналы с шумом отказов и шумом помех в единый канал с шумом

(рис. 77). Используя полученные ранее соотношения, можно определить количество информации, передаваемой через такой канал с шумом:

$$I(X_0, Y_0) = H(Y_0) \quad H(Y_0/X_0),$$

где $H(Y_0)$ — энтропия на выходе канала, $H(Y_0/X_0)$ — условная энтропия приема множества сообщений $Y_0$ при условии, что задаваемое на входе канала множество сообщений $X_0$ известно.

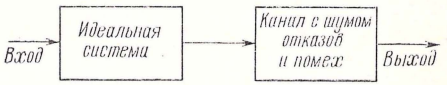


Рис. 77.

Было получено, что потери информации для случая равновероятных сообщений дискретного характера при использовании кодов, рассчитанных на исправление ошибок ($d=2s+1$):

$$H(Y_0/X_0) = -P_0 \log_2 P_0 - (1-P_0)\log_2(1-P_0) + \\ + P_0 \log_2 (M-1),$$

где $P_0$ — вероятность ошибки при выявлении сообщения, возникающей под действием помех в линии связи и отказов технических средств, в данном случае $P_0 = P_{тр}$.

Рассмотрим некоторые примеры оценки потерь. Если положить, что надежность является идеальной, то потери информации характеризуются как следствие воздействия помех в линии связи и можно выполнить анализ помехоустойчивости на основе принятой модели помех. Если на входе канала имеется некоторое множество символов $X$, а на выходе — множество символов $Y$, то потери информации

$$H(Y/X) = -p \log_2 p - (1-p)\log_2(1-p),$$

где $p$ — вероятность искажения символа в двоичном симметричном канале с шумом. Величина $p$ устанавливается исходя из уровня помех в канале связи и реализации первой решающей схемы.

Применим этот подход для оценки надежности. Рассмотрим некоторое кодирующее устройство, которое преобразует множество сообщений $X_0$ во множество символов $X$. Вследствие его ненадежности получаем некоторое промежуточное множество символов $Z$, которые могли бы быть сформированы идеальным кодирующим устройством без неисправностей, т. е. представим кодирующее устройство в виде последовательного соединения идеального кодирующего устройства, преобразующего множество сообщений $X_0$ в множество символов $Z$, и некоторого канала с шумом отказов, где осуществляется преобразование символов множества $Z$ в символы множества $X$, которые поступают

далее в канал связи с помехами. При идеальном процессе кодирования и идеально надежном кодирующем устройстве энтропия сообщения $H(X_0)=mH(X)$, где $m$ — число информационных символов в коде сообщения, $H(X)$ — энтропия источника.

Из-за ненадежности кодирующего устройства возникают потери информации, численно равные $H(X_0)-mH(X)$. Учитывая неисправности и их последствия в виде отказов, статистические свойства кодирующего устройства можно описать следующей совокупностью вероятностей:

$$p(x_i/X_0)=\sum_{j=1}^{k} p(x_i/z_j)\,p(z_j/X_0),\ \text{где}\ x_i\in X,\ z_j\in Z.$$

Потери информации при ненадежном кодировании оценим как потери при переходе от множества $Z$ к множеству $X$. Тогда

$$H(X/Z)=-\sum_{i=1}^{K}\sum_{j=1}^{K} p(z_j)\,p(x_i/z_j)\log_2 p(x_i/z_j).$$

Возникает задача определения условных вероятностей преобразования символов множества $Z$ в символы множества $X$. Рассмотрим эти вероятности для случая, когда формируется код, последовательный во времени, и отказы, возникающие в аппаратуре кодирующего устройства, подчиняются во времени закону распределения Пуассона.

В большинстве кодирующих устройств удается выделить индивидуальные цепи формирования каждого элемента кода с интенсивностью отказа цепи $\lambda_1$. Тогда вероятность правильного прохождения символа по выделенной цепи формирования можно оценить как $p(x_i/z_j)=\exp(-\lambda_1 t)$ при $i=j$. Соответственно вероятность ошибки при формировании одного элемента кода в кодирующем устройстве под действием отказов равна $p(x_i/z_j)=1-\exp(-\lambda_1 t)$ при $i\neq j$. В отдельных случаях можно считать, что наиболее вероятна трансформация одного символа в другой при наличии одиночной ошибки в каждой из двух соседних цепей формирования элемента кода. Тогда $p(x_i/z_j)=\lambda_1^2 t^2\exp(-2\lambda_1 t)$. Здесь использованы первые члены распределения Пуассона, полученные из соотношения $p(i)=[(\lambda_1 t)^i/i!]\exp(-\lambda_1 t)$, где $p(i)$ — вероятность возникновения $i$ отказов за время $t$. Если учесть возникновение трансформации от отказов большей кратности, то получаем

$$p(x_i/z_j)=p(mp/2)\,p(2)+p(mp/3)\,p(3)+\ \ldots$$
$$\ldots+p(mp/i)\,p(i)+\ \ldots+p(mp/n)\,p(n)$$

при $j\neq i$, $j\neq 1$, $i\neq 1$, где $p(mp/i)$ — вероятность трансформации в случае возникновения ошибки кратности $i$, $p(i)$ — вероятность возникновения отказа кратности $i$ (в $i$ цепях формирования элемента кода). Если формируется двоичный код с равновероятным появлением каждого символа, то мож-

но использовать модель двоичного симметричного канала, для которого

$$p(x_1/z_1) = p(x_2/z_2) = \exp(-\lambda_1 t),\ p(x_2/z_1) = p(x_1/z_2).$$

Потери информации могут быть определены в этом случае через условную энтропию

$$H(X/Z) = 2[-p(z_1)\,p(x_1/z_1)\log_2 p(x_1/z_1) -$$
$$-p(z_1)\,p(x_2/z_1)\log_2 p(x_2/z_1)] = -\exp(-\lambda_1 t)\log_2\exp(-\lambda_1 t) -$$
$$-[1-\exp(-\lambda_1 t)]\log_2[1-\exp(-\lambda_1 t)].$$

Нетрудно видеть, что таким способом можно оценить потери, вызванные действием неисправности аппаратуры кодирующего устройства при формировании символов $x$, поступающих далее в канал связи.

Если $\lambda_1$ — интенсивность отказа одной цепи выявления элемента кода, то полученные соотношения могут быть использованы для оценки потерь информации, возникающих вследствие ненадежности декодирующего устройства. Представляет интерес объединить все потери информации в единую совокупность и получить суммарные потери, вызванные действием сбоев, отказов как в кодирующем, так и в декодирующем устройствах. В этом случае $\lambda_1$ — суммарная интенсивность отказа цепи формирования и цепи выявления одного элемента кода. Граф условного канала с шумом отказов представлен на рис. 78: $z_1$, $z_2$ — символы, поступающие на вход канала, $x_1$, $x_2$ — символы на выходе канала с шумом отказов. Как видно из графа, имеем модель двоичного симметричного канала. Следовательно, физически значение энтропии $H(X/Z)$ характеризует рассеяние информации в условном канале с шумом отказов на один символ формируемого кода. Так как символы $x$ под действием помех подвергаются искажению в реальном канале связи, то можно представить себе граф некоторого условного канала с шумом отказов и помех, состоящий из графа канала, характеризующего шум отказов, и графа канала с шумом помех. Возникает единый условный канал с шумом отказов и помех, характеризующийся вероятностью искажения символа:

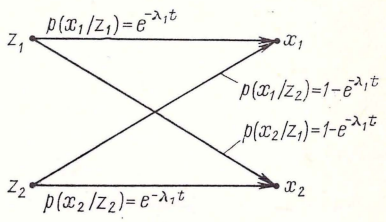


Рис. 78.

$$p = p_{\text{пом}}(1 - p_{\text{отк}}) + p_{\text{отк}}(1 - p_{\text{пом}}) + p_{\text{отк}}p_{\text{пом}},$$

где $p_{\text{пом}}$ — вероятность искажения символа под действием помех в линии связи, $p_{\text{отк}}$ — вероятность искажения символа

по причине отказов аппаратуры. Для реальных условий передачи информации в АСУ $p_{\text{пом}} \ll 1$, $p_{\text{отк}} \ll 1$, а поэтому приближенно можно считать, что $p = p_{\text{пом}} + p_{\text{отк}}$.

Окончательно потери информации при переходе символов множества $Z$ к символам множества $X$ и далее символов множества $X$ к символам множества $Y$ можно оценить энтропией

$$H(Y/Z) = -p \log_2 p - (1-p) \log_2 (1-p).$$

Численный расчет потерь информации зависит от конкретных моделей помех и отказов аппаратуры системы, а также от вида избыточности при передаче и преобразовании информации в ПСПИ.

Представляет интерес оценка той минимальной избыточности, которую необходимо внести, чтобы компенсировать потери информации. Рассмотрим потери информации, возникающие по причине ненадежности аппаратуры. Обозначим эти потери индексом «а», а потери информации вследствие искажения в канале связи — индексом «и». Для устранения потерь, связанных с ненадежностью кодирующего устройства, введем в него избыточность. Тогда вместо $m$ цепей возникает $n$ цепей формирования элементов кода, т. е. вводится $k$ контрольных цепей. Потребуем, чтобы под действием отказов искажалось не более $k$ контрольных элементов в коде, тогда $k = H_a(X/Z)\, n$. Так как $k = n - m$, то $nH_a(H/Z) = n - m$.

Вводя коэффициент избыточности в аппаратуре $\gamma_a = n/m$, получаем, что

$$\gamma_{a\,\min} = 1/[1 - H_a(X/Z)].$$

Можно показать, что в реальных условиях доля необходимой вводимой избыточности обычно оказывается больше данной величины, которая в определенной степени указывает минимально необходимую избыточность. Величина $\gamma_a$ является функцией от параметра $\lambda_1 t$: с увеличением $\lambda_1 t$ значение $\gamma_a$ увеличивается. Таким образом, при заданных вероятностных показателях качества функционирования кодирующего устройства с увеличением времени эксплуатации исходную избыточность необходимо увеличивать.

Аналогичные результаты могут быть получены, если рассматривать условный канал с шумом помех. Тогда получаем коэффициент избыточности в информации

$$\gamma_{и\,\min} = 1/[1 - H_и(Y/Z)],$$

где $H_и(Y/Z)$ — потери информации в канале связи, вызванные действием помехи.

Используя данный информационный подход к анализу потерь, вызванных как действием помех, так и действием

отказов аппаратуры, можно ввести коэффициент избыточности $\gamma$ в информации и в аппаратуре и определить его нижний уровень в виде

$$\gamma_{\min} = 1/[1 - H(Y/X)],$$

где $H(Y/X)$ — потери информации, вызванные действием помех и отказов аппаратуры.

Вероятностным критерием, оценивающим потери информации, как указывалось выше, является вероятность потерь. Если потери вызваны действием помех в канале связи и отказами аппаратуры, то данная вероятность $P_{\text{пот.н}} = 1 - P_{\text{пр}}$. Однако для получения общей вероятности потерь в системе необходимо учесть еще потери, вызванные отказом в обслуживании.

Приемная часть подсистемы сбора и передачи информации АСУ функционирует как однолинейная или многолинейная обслуживающая система, на вход которой подается совокупность информационных потоков. Потоки сообщений могут быть рассмотрены со стороны приемной части ПСПИ как потоки заявок, а сама приемная часть является системой обслуживания. Обслуживание в этом случае может быть формализовано в двух вариантах.

1. Обслуживание до выявления кодовой комбинации на приемной стороне, когда число передач зависит от действующих помех в канале связи, в том числе, когда принятая кодовая комбинация отождествляется с одной из заданных, прекращается запрос по обратному каналу и обслуживание данного сообщения заканчивается. Время обслуживания одного сообщения является переменным, т. е. получаем обслуживающую систему с переменным временем обслуживания, зависящим от уровня помех в момент передачи информации в канале связи.

2. Обслуживание осуществляется по одной передаче в том случае, если допускается многократная выдача одного и того же сообщения. Повторные передачи рассматриваются как индивидуальные заявки. Тогда время обслуживания каждого сообщения постоянно. Независимо от исхода обслуживания данной заявки последующая передача сообщения рассматривается как новая заявка. Формализацией такой системы является обслуживающая система с постоянным временем обслуживания.

Расчет вероятности потерь по причине отказа в обслуживании можно вести исходя из известных в теории массового обслуживания соотношений, связывающих параметры входного потока и параметры обслуживающей системы, такие, как интенсивность потока заявок $\lambda$, интенсивность обслуживания $\mu$, число обслуживающих приборов $N$, длина очереди (объем запоминающих устройств перед обслуживающими приборами) $l$. При этом существенное влияние оказывает закон распределения времени обслуживания, а также организация очереди,

которая может быть выбрана раздельной, когда перед каждым декодирующим устройством устанавливается свое индивидуальное запоминающее устройство, либо общей, когда очередь организуется в запоминающем устройстве, устанавливаемом перед всеми декодирующими устройствами. Особенностью оценки вероятности отказа в обслуживании является то, что увеличение избыточности в передаваемой информации увеличивает время обслуживания каждого сообщения. В соответствии с этим снижается интенсивность обслуживания, а поэтому наблюдается такая закономерность: с увеличением избыточности в информации при отсутствии обратной связи в передаче ухудшаются характеристики обслуживания из-за снижения интенсивности обслуживания. Следовательно, вероятность отказа в обслуживании необходимо рассчитывать с учетом всех видов избыточности, вводимой в информацию.

Чтобы оценить потери, связанные с отказом в обслуживании, рассмотрим виды избыточности, которые могут вводиться в передаваемую информацию и в аппаратуру системы для обеспечения заданного качества функционирования технических средств АСУ.

## § 3. Избыточность как средство обеспечения заданного качества функционирования технических средств АСУ

Одним из наиболее эффективных методов обеспечения требуемого качества функционирования комплекса технических средств АСУ в целом и его отдельных подсистем является введение избыточности. Для обеспечения заданной верности преобразования информации с учетом помехоустойчивости передачи и надежности ее обработки необходимо вводить избыточность как в передаваемый сигнал, так и в аппаратуру системы. Поэтому возникает несколько разновидностей избыточности, которые используются в настоящее время при практической реализации информационных систем.

Под избыточностью будем понимать дополнительный объем передаваемого сигнала или дополнительный объем аппаратуры, которая используется для обеспечения заданных качеств функционирования технических средств. Избыточность передаваемого сигнала, направленную на повышение помехоустойчивости передачи, назовем *информационной* избыточностью. Избыточность в аппаратуре системы, которая способствует обеспечению заданной работоспособности, определим как *аппаратурную* избыточность.

Каждый вид информационной или аппаратурной избыточности может реализоваться на основе некоторого кода, в который вводится избыточная часть. Тогда возникает *кодовая* избыточность, определяющая использование кода, способного либо обнаруживать, либо обнаруживать и исправлять ошибки.

Для повышения качества функционирования технических средств можно использовать многократное повторение однотипных операций, блоков, сигналов и т. д. для получения требуемой помехоустойчивости или надежности. Такую избыточность назовем избыточностью повторения.

Для ПСПИ характерны следующие виды кодовой избыточности: информационная, аппаратурная, общая.

Ранее показано, что при кодовой информационной избыточности передаваемый по линии связи сигнал представляется некоторым избыточным кодом, в котором имеется определенная доля контрольных элементов, и такой код можно охарактеризовать коэффициентом избыточности $\gamma_и$. Если код неизбыточен, то $m=n$ и $\gamma_и=1$. С введением избыточности $\gamma_и>1$ и определяется числом передаваемых сообщений и методом построения кода. Вместе с тем $\gamma_и$ существенно зависит от корректирующей способности кода: возрастает, например, для $M=64$ от $\gamma_и=1{,}5$ (для кода $s=1$) до $\gamma_и=2{,}5$ (для кода $s=3$) при основании кода $K=2$. Кодовая информационная избыточность является наиболее распространенным видом избыточности в реальных системах передачи информации в АСУ, где наибольшее распространение получили циклические коды. Характерно, что с увеличением длины кода относительная доля контрольных элементов в коде уменьшается и соответственно уменьшается коэффициент $\gamma_и$ для той же корректирующей способности кода.

Кодовая аппаратурная избыточность предполагает реализацию технических средств, функционирующих в избыточных кодах с коэффициентом избыточности $\gamma_а$, что означает появление в системе или устройстве разрешенных и запрещенных состояний, в связи с чем возникает возможность исправления ошибок в выходных сигналах, являющихся следствием отказов аппаратуры системы. Чтобы реализовать кодовую аппаратурную избыточность, необходимо ввести ряд дополнительных элементов, обладающих конечной надежностью и подверженных также действию отказов и сбоев. Это ограничивает возможности использования кодовой аппаратурной избыточности в реальных устройствах, особенно при использовании кодов с высокой корректирующей способностью, а поэтому существуют границы допустимого объема дополнительно вводимой аппаратуры, до которых кодовая аппаратурная избыточность является эффективной.

Весьма интересной разновидностью кодовой избыточности является общая кодовая избыточность. В этом случае исправляющая способность кода определяется исходя из суммарного действия помех в линии связи и отказов технических средств в системе. Выявление ошибок осуществляется декодирующим устройством, которое располагается на приемной части ПСПИ, и при соответствии корректирующей способности кода интенсивности помехи и интенсивности отказов возникает

возможность нормального функционирования системы даже при наличии отказавших элементов и помех в линии связи. Таким образом, общая кодовая избыточность является совместным средством обеспечения качества функционирования ПСПИ с учетом действия помех и отказов технических средств.

Перейдем к рассмотрению избыточности повторения, которая может реализоваться в виде информационной, аппаратурной и общей избыточности повторения.

Информационная избыточность повторения предполагает повторение информации в канале связи во времени и по частоте, или повторение по временным и частотным каналам. Использование частотных каналов при ограниченной общей полосе пропускания канала связи означает увеличение длительности элементарных посылок в каждом отдельном канале. Поэтому избыточность повторения по частоте вызывает увеличение длительности передачи информации и неразрывно связана с временной избыточностью. Практически избыточность повторения по времени означает многократную передачу одного и того же сообщения при различных алгоритмах управления передачей, постоянном или переменном числе передач одного и того же сообщения. В последнем случае избыточность становится переменной и необходимы дополнительные сигналы, передаваемые по обратному каналу, которые прекращают последующую выдачу одного и того же сообщения.

В чистом виде информационная избыточность повторения означает использование для передачи каждого сообщения неизбыточного кода. Коэффициент избыточности $\gamma = a$, где $a$ — число передач одного и того же сообщения. В случае реализации системы с повторением на заданное число передач $a$ зависит от требуемой верности передаваемой информации и некоторой расчетной, выбранной заранее интенсивности помех в канале связи. Обычно при фиксированном $a$ система рассчитывается исходя из наихудших условий передачи информации, а поэтому в ряде случаев постоянная информационная избыточность повторения оказывается крайне неэффективной, так как возможны случаи передачи с высокой избыточностью при малых интенсивностях помехи.

Следует отметить, что использование информационной избыточности повторения (учитывая, что код каждого сообщения является неизбыточным) требует применения на приемной стороне устройств накопления, осуществляющих мажоритарный способ выявления неизбыточных кодов и предполагающих выявление сообщений при возникновении на приемной стороне не менее $a_н$ раз одного и того же кода из $a$ передач, где $a_н$ — число накоплений. Избыточность повторения по частоте означает введение в систему параллельных частотных каналов. В этом случае накопительное устройство на приемной стороне реализуется путем пороговой схемы и необходимость запоминания ранее

принятых сообщений отсутствует. Коэффициент избыточности $\gamma = k_f$, где $k_f$ — число параллельных частотных каналов, в каждом из которых передается избыточный код данного сообщения.

Аппаратурная избыточность повторения способствует повышению надежности системы. Если она реализуется в виде повторения отдельных элементов, блоков или устройств, то получаем резервирование с использованием мажоритарных органов для обнаружения и исправления ошибок отказов. Если же эта избыточность сопровождается повторением функционирования устройства во времени, то она позволяет исключить последствия сбоев аппаратуры.

Общая избыточность повторения означает параллельное введение избыточности как в передаваемый сигнал, так и в структуру системы. В этом случае повторяемые во времени элементы сообщения соответствуют повторяемым в аппаратуре конструктивным элементам технических средств. Возникают резервные цепи формирования и выявления элементов кода на передающей и приемной частях ПСПИ. При этом появляется возможность за счет повторения обнаруживать, а также обнаруживать и исправлять ошибки, возникающие под действием отказов конструктивных элементов системы и вследствие наличия помех в линии связи. Следует отметить, однако, что использование общей избыточности повторения в чистом виде вызывает значительное снижение скорости передачи информации и увеличение числа конструктивных элементов в системе.

В практических случаях реализации передачи и обработки информации наиболее эффективным оказывается совместное использование различных видов информационной и аппаратурной избыточности. Прежде всего весьма полезно совмещение информационной кодовой и информационной избыточности повторения, когда используется повторение кодов, обладающих определенной избыточностью. В этом случае введение в код избыточности с целью обнаружения ошибок в сочетании с повторением позволяет получить наиболее эффективные средства обеспечения помехоустойчивости передачи при сравнительно малых аппаратурных затратах. Вместе с тем использование сочетаний тех же видов избыточности в аппаратуре позволяет при малом дополнительно вводимом объеме аппаратуры обнаруживать и исправлять последствия отказов или сбоев.

Поскольку АСУ является человеко-машинной системой, то человек может исправить последствия неисправностей. Поэтому при реализации аппаратурной избыточности можно ограничиться случаем обнаружения ошибок, а функцию исправления возложить на человека. Для этого в технических средствах необходимо предусмотреть устройство контроля, которое на основе тестового контроля или контроля по рабочим выходным

сигналам позволяет выдать информацию о возникновении неисправности.

Возложив функцию исправления последствий неисправностей на аппаратуру, получим функционально безотказную систему, и до определенного момента времени, если не учитывать возможность восстановления, мы можем не выявить наличия отказа в системе. Однако при автоматическом исправлении ошибок отказа аппаратура усложняется за счет введения специальных корректирующих устройств, усложняется логика функционирования технических средств. Весьма существенны поэтому тип, а также уровень вводимой избыточности.

Поэтому при проектировании комплекса технических средств возникает ряд задач, связанных с оптимальным выбором методов и средств, реализующих избыточность.

Большое значение имеет выбор коэффициента избыточности. Прежде всего полезно введение избыточности с переменным коэффициентом $\gamma$, т. е. переход к системе с адаптацией. Применительно к информационной избыточности адаптация может быть реализована либо в виде систем с обратной связью, либо в виде специальных адаптивных систем, в которых учитывается состояние канала и система приспосабливается к каждому состоянию канала, характеризующемуся своими параметрами действующих ошибок.

При использовании переменной информационной кодовой избыточности получаем систему с адаптивным кодированием, в которой переменным является соотношение между числом информационных и контрольных элементов в коде. При использовании информационной избыточности повторения за счет наличия цепи обратной связи осуществляем переменное число передач, т. е. система подстраивается к условиям действия помех в момент передачи информации. Для аппаратурной избыточности при переменном коэффициенте $\gamma$ можно построить адаптивный мажоритарный орган. Меняя порог мажоритарного органа, удается реализовать оптимальные варианты коррекции ошибок отказов аппаратуры. Однако даже в случае введения переменной избыточности объем аппаратуры системы является постоянным, выбираемым из наихудших условий функционирования системы. Таким образом, основной целью введения избыточности является снижение потерь информации, возникающих в системе.

Потери информации зависят от уровня введенной избыточности, а поэтому необходимо охарактеризовать избыточность, введенную в систему, количественно. Для оценки избыточности можно использовать информационный подход и введенное ранее понятие кодового расстояния. Учитывая наличие различных видов избыточности, введем понятие информационного кодового расстояния, которое характеризует число переходов между кодовыми комбинациями, отображающими сигнал в линии свя-

зи, и обозначим его $d_{\text{и}}$, причем $d_{\text{и}} \geqslant r_{\text{и}} + s_{\text{и}} + 1$. Для информационной избыточности можно получить связь между коэффициентом избыточности и числом исправляемых ошибок исходя из найденных информационных пределов избыточности. Тогда для двоичного кода, исправляющего $s_{\text{и}}$ ошибок, при $r_{\text{и}} = s_{\text{и}}$

$$M = 2^m \leqslant 2^{\gamma_{\text{и}} m} \Big/ \sum_{i=0}^{s_{\text{и}}} C^i_{\gamma_{\text{и}} m},$$

соответственно для $K$-ичного кода

$$M = K^m \leqslant K^{\gamma_{\text{и}} m} \Big/ \sum_{i=0}^{s_{\text{и}}} C^i_{\gamma_{\text{и}} m} (K-1)^i.$$

Информационную избыточность повторения количественно можно также оценивать кодовым расстоянием $d_{\text{и}}$, при этом $d_{\text{и}} = a$, если повторяется код с кодовым расстоянием $d = 1$. Вводя на приемной стороне $a_{\text{н}}$ накоплений, можно обнаружить $r_{\text{и}}$ ошибок, причем $r_{\text{и}} = a_{\text{н}} - 1$, так как при числе ошибок, равном числу накоплений, происходит трансформация сообщения. Учитывая, что $d_{\text{и}} \geqslant r_{\text{и}} + s_{\text{и}} + 1$, получаем $a \geqslant a_{\text{н}} - 1 + s_{\text{и}} + 1$, откуда $s_{\text{и}} \leqslant a - a_{\text{н}}$. Таким образом, при использовании информационной избыточности повторения и накопления на приемной стороне удается исправить $s_{\text{и}}$ ошибок. Следует отметить, что это справедливо, если $a - a_{\text{н}} < a_{\text{н}}$.

При совместном использовании кодовой информационной и информационной избыточности повторения обычно осуществляется $a$-кратная передача кода, обнаруживающего ошибки, с кодовым расстоянием $d_{\text{о}}$. Поэтому кодовое расстояние

$$d_{\text{и}} = a d_{\text{о}}, \quad a d_{\text{о}} \geqslant r_{\text{и}} + s_{\text{и}} + 1.$$

Используя информационный подход, можно осуществить количественную оценку аппаратурной избыточности. Воспользуемся геометрическим представлением кодов: пусть каждому состоянию системы соответствует некоторая точка в пространстве. Тогда можно выделить точки, отвечающие исправным и неисправным состояниям системы. Если вокруг каждого разрешенного состояния, так же как в теории помехоустойчивого кодирования, образовать некоторую область и считать, что она включает в себя ряд разрешенных состояний, то при попадании системы в некоторое неисправное состояние $V_i'$ данное состояние может быть отнесено к заданному состоянию $V_i$, если между $V'$ и $V_i$ имеется число переходов $s_{\text{а}}$, где $s_{\text{а}}$ — число ошибок отказа, исправляемых за счет введения избыточности. Соответственно между центрами разрешенных областей должно быть аппаратурное расстояние $d_{\text{а}}$, причем $d_{\text{а}} \geqslant r_{\text{а}} + s_{\text{а}} + 1$, где $r_{\text{а}}$ — число ошибок отказа, обнаруживаемых в техническом средстве. Если в системе не предусматривается каких-либо способов введения аппаратурной избыточности, то $d_{\text{а}}$ в большинстве случаев равно единице, и система является неизбыточной.

При введении аппаратурной избыточности увеличение аппаратурного кодового расстояния достигается за счет охвата избыточностью какой-то части аппаратуры. При этом появляются элементы, которые осуществляют коррекцию ошибок и сами избыточностью не охватываются. Такие элементы не исправляют и не обнаруживают собственных ошибок сбоев и отказов, а поэтому можно считать, что при введении аппаратурной избыточности значение $d_a$ становится больше единицы лишь для какой-то части системы с избыточностью и $d_a=1$ для остальной группы элементов.

Оценим численно значение $d_a$. Для кодовой аппаратурной избыточности значение $d_a$ связано с коэффициентом аппаратурной избыточности соотношениями, определяемыми приведенными выше информационными пределами избыточности. Для аппаратурной избыточности повторения кодовое расстояние $d_a=\alpha+1$, где $\alpha$ — число избыточных повторяющих каналов. В этом случае наряду с устройством, охватываемым избыточностью, появляется $\alpha$ дополнительных параллельных устройств, осуществляющих резервирование. Однако возникают и неизбыточные элементы, к которым можно отнести мажоритарные узлы, объединяющие параллельные аппаратурные каналы. Эти узлы не исправляют собственных ошибок сбоев и отказов. Так как для исправления ошибок необходимо иметь $d_a\geqslant 3$, то $\alpha\geqslant 2$. Реализация аппаратурной избыточности повторения вызывает значительное усложнение аппаратуры системы. Поэтому на практике используется совместно кодовая аппаратурная избыточность и аппаратурная избыточность повторения.

Количественной оценкой совместного использования этих видов избыточности может служить кодовое расстояние $d_a$, определяющееся с учетом числа переходов кода $d_o$, в котором функционирует техническое устройство, и числа параллельных избыточных повторяющих каналов $\alpha$: $d_a=d_o(\alpha+1)$. Опыт использования аппаратурной избыточности показывает, что наиболее простая схемная реализация возникает при $d_o=2$, т. е. когда устройство функционирует в коде, обнаруживающем ошибку. Тогда уже при $\alpha=1$ имеем $d_a=4$, и появляется возможность обнаружения двойной и исправления одиночной ошибок отказа. Если увеличить $d_o$ до трех, то получаем устройство, функционирующее в корректирующем коде, и с учетом аппаратурных избыточных повторяющих каналов, т. е. $\alpha$-кратного повторения, возникает возможность исправления ряда ошибок отказов.

Для указанных выше общих методов введения избыточности кодовое расстояние $d=d_a=d_и$. Удается исправить $s$ ошибок для $d\geqslant 2s+1$. При этом в $s$ должны входить все возможные ошибки, связанные с действием помех в каналах связи, и ошибки, являющиеся следствием неисправностей аппаратуры.

Использование общей избыточности иллюстрируется рис. 79, где представлен граф прохождения кодовой комбинации $x_{0j}$, выбираемой из входного множества сообщений $X_0$. В этом графе горизонтальный путь соответствует прохождению кодовых комбинаций при отсутствии ошибок. Остальные пути ведут к правильному выявлению, если суммарное отклонение от горизонтальной оси не превышает числа исправляемых ошибок $s$. Выделены области исправления ошибок, которые соответствуют максимально допустимым исправляемым искажениям. Если код имеет одинаковые возможности по исправлению ошибок отказов на передающей и на приемной сторонах и соответственно в каналах связи, то $s_{\text{а пер}}=s_{\text{и}}=s_{\text{а пр}}$. Случаю использования всех возможных корректирующих свойств кода соответствует множество различных путей; однако если возникает хотя бы одно отклонение, соответствующее радиусу указанных окружностей, то это означает полное использование корректирующей способности кода, и далее правильный прием возможен, если во всех остальных трактах передачи при формировании и выявлении информации не будет иметь место ошибка.

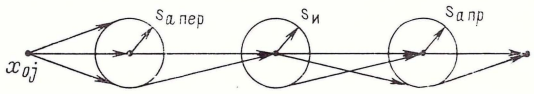


Рис. 79.

Таким образом, при общей кодовой избыточности используются известные способы кодирования, которые приложимы к введению избыточности в передаваемую информацию и в аппаратуру технических средств.

Общую избыточность повторения можно охарактеризовать кодовым расстоянием $d=\alpha+1$, где $\alpha$ — число повторяющих каналов, причем число как повторений информации, так и повторяющих каналов в аппаратуре. Очевидно, что $\alpha>1$. Даже для исправления одиночных ошибок необходимо иметь $\alpha=2$, что вызывает усложнение аппаратуры, а поэтому представляется полезным использовать совместно общую кодовую и общую избыточность повторения. Тогда используется основной код с числом переходов $d_0$ и применяется $\alpha$ параллельных каналов: $d=d_0(\alpha+1)$. Отсюда уже при $\alpha=1$ и $d_0=2$ обеспечивается исправление одиночной ошибки, вызванной действием помехи или неисправностью технических средств.

Если исходные данные по действующим помехам в канале связи и неисправностям аппаратуры требуют различного уровня введения аппаратурной и информационной избыточности, то можно использовать соответствующие значения $\alpha$ и $d_0$ или изменить корректирующие свойства кодов. При этом удается сместить количественные показатели избыточности либо

в сторону увеличения $d_а$, либо в сторону возрастания $d_и$ в зависимости от конкретных условий эксплуатации технических средств.

Поскольку для получения одних и тех же показателей качества функционирования технических средств могут быть использованы различные виды избыточности, то представляется целесообразным найти границы применения тех или иных видов избыточности при создании технических средств передачи и обработки информации. Следует отметить, что получаемые при этом пределы избыточности в значительной степени зависят от ограничений, накладываемых на возмущающие воздействия в системе и соответственно на модели, которые отображают исходные процессы случайных воздействий.

Если предположить, что надежность технических средств является идеальной, то, зная модель помех или ошибок, возникающих в канале связи, можно найти границы эффективного использования информационной избыточности при следующем подходе. За исходный примем код, обнаруживающий ошибку, для которого $r_и=1$, $s_и=0$, $d_и=2$, и путем введения дополнительных контрольных элементов в коде перейдем к коду, исправляющему $s_и$ ошибок, с числом элементов $\gamma_и m$. Тогда можно считать, что корректирующий код может быть использован даже в том случае, если он дает такую же вероятность правильного приема, как и обнаруживающий. Рассмотрим это на примере некоторой модели ошибок. Предположим, что ошибки в канале связи подчиняются закону распределения Пуассона, т. е. вероятность попадания $i$ ошибок на код длиной $n$

$$p(i) = [(\gamma_s m \tau f_{со})^i / i!] \exp(-\gamma_s m \tau f_{со}).$$

Для обнаруживающего кода вероятность прохождения (правильного приема) сообщения есть вероятность непопадания ошибок на код, т. е.

$$P_{пр} = p(0) = \exp(-\gamma_0 m \tau f_{со}),$$

где $\gamma_0$ — коэффициент избыточности обнаруживающего кода. Соответственно для кода, исправляющего $s_и$ ошибок, вероятность правильного приема сообщения есть вероятность того, что на код попадает не более чем $s_и$ ошибок. Поэтому можно записать, что

$$P_{пр\,s} = \sum_{i=0}^{s_и} p(i) = \sum_{i=0}^{s_и} [(\gamma_s m \tau f_{со})^i / i!] \exp(-\gamma_s m \tau f_{со}),$$

где $\gamma_s$ — коэффициент избыточности кода, исправляющего $s_и$ ошибок.

Эффективность использования избыточного кода, исправляющего ошибки, оценим по эквивалентному снижению средней частоты ошибок помехи $f_{со}$, т. е. будем считать, что

введение коррекции снижает среднюю частоту ошибок $f_{со}$ по исходному коду до величины $f_{со.к}$. Тогда вероятность $P_{пр\,s}$ можно рассматривать как вероятность прохождения сообщения при использовании кода, обнаруживающего ошибку, если в канале связи возникают ошибки с приведенной средней частотой $f_{со.к}$. Тогда

$$\exp(-\gamma_0 m\tau f_{со.к}) = \sum_{i=0}^{s_и} [(\gamma_s m\tau f_{со})^i/i!] \exp(-\gamma_s m\tau f_{со}).$$

Отсюда находим

$$f_{со.к} = f_{со}\gamma_s/\gamma_0 \left[1 - 1/\sum_{i=0}^{s_и} (\gamma_и m\tau f_{со})^i/i!\right].$$

Ниже представлены предельные значения $\gamma_s/\gamma_0$, при которых допустимо использование кода в режиме исправления одной или двух ошибок.

| Интенсивность ошибок . . . . | 0 | 0,4 | 0,8 | 1,2 | 1,6 | 2,0 |
|---|---|---|---|---|---|---|
| $s = 1$, $\gamma_s/\gamma_0$ . . . | любое | 3,4 | 2,2 | 1,2 | 1,6 | 1,5 |
| $s = 2$, $\gamma_s/\gamma_0$ . . . | любое | 20 | 6,7 | 4 | 3 | 2,5 |

Поскольку задана минимально допустимая вероятность прохождения $[P_{пр}]_{доп}$, то следует обеспечить, чтобы

$$\exp(-\gamma_0 m\tau f_{со.к}) \geqslant [P_{пр}]_{доп},$$

тогда необходимый уровень снижения средней частоты ошибок устанавливается из условия

$$f_{со.к}/f_{со} \leqslant -1/\gamma_0 m\tau f_{со} \ln [P_{пр}]_{доп}.$$

Исходя из этого, можно выбрать такое значение $\gamma_s/\gamma_0$, которое обеспечит заданное отношение $f_{со.к}/f_{со}$. Аналогичным способом можно найти и границы использования информационной избыточности повторения.

Положим, что код, обнаруживающий ошибку, передается $a$ раз и на приемной стороне отсутствует накопление, т. е. память по принимаемым сообщениям. Вероятность прохождения при однократной передаче для пуассоновской модели ошибок в канале связи определяется, как ранее было получено, в виде

$$P_{пр} = \exp(-\gamma_0 m\tau f_{со}).$$

Соответственно вероятность ошибки при однократной передаче $P_0 = 1 - P_{пр}$. Если ошибки подчиняются закону Пуассона, то последовательные передачи сообщений можно рассматривать как независимые события, и вероятность прохождения сообщения за $a$ передач

$$P_{пр\,a} = 1 - P_0^a = 1 - [1 - \exp(-\gamma_0 m\tau f_{со})]^a.$$

Введем, как и ранее, приведенное значение средней частоты ошибок $f_{co\,a}$, учитывая, что вероятность прохождения $P_{пр\,a}$ может быть обеспечена исходным обнаруживающим кодом, но при наличии ошибок с приведенной средней частотой $f_{co\,a}$, тогда

$$\exp(-\gamma_0 m\tau f_{co\,a}) = 1 - [1 - \exp(-\gamma_0 m\tau f_{co})]^a,$$

$$f_{co\,a} = a f_{co} \exp(-\gamma_0 m\tau f_{co}) [1 - \exp(-\gamma_0 m\tau f_{co})]^{a-1} \{1 - [1 - \exp(-\gamma_0 m\tau f_{co})]^a\}^{-1}.$$

Граничные значения $\gamma_s/\gamma_0$, до которых код, исправляющий одну ошибку, дает большую вероятность прохождения, чем двукратно переданный код, обнаруживающий одну ошибку, приведены ниже.

| Интенсивность ошибок . . . | 0 | 0,4 | 0,8 | 1,2 | 1,6 | 2,0 |
|---|---|---|---|---|---|---|
| $\gamma_s/\gamma_0$ . . . . . | любое | 1,7 | 1,55 | 1,5 | 1,45 | 1,4 |

Аналогично могут быть рассмотрены и другие варианты совместного использования различных видов информационной избыточности. Следует отметить, что большинство каналов передачи информации в АСУ, особенно каналов, соединяющих узлы коммутации сообщений, отличается нестационарностью, несимметричностью, в них проявляется зависимость искажений последующих символов от предыдущих. Для модели независимых ошибок в канале связи можно получить вероятности ошибки, которые на два порядка отличаются от действительных вероятностей, рассчитанных по конкретным моделям ошибок, выявленных статистическим анализом канала связи. В настоящее время широко применяются исследования передачи информации по каналам, в которых ошибки группируются в пакеты.

Под пакетом обычно понимают некоторую последовательность символов, которые ограничены искаженными символами и характеризуются некоторой плотностью ошибочных символов $q_0$. В [9] предлагается формула для расчета вероятности ошибки при передаче сообщений кодом длиной $n_0$ при условии, что пакеты ошибок возникают независимо друг от друга и расстояние между двумя ошибочными символами не превышает некоторой величины

$$P_0 = \frac{p}{q_0} \sum_{l=1}^{l_{max}} \left(1 + \frac{n_0 - 1}{l}\right) \frac{l p_l}{\sum_{l=1}^{l_{max}} l p_l},$$

где $l$ — длина пакета, $p_l$ — условная вероятность появления пакета ошибок длиной $l$, $p$ — вероятность искажения символа в канале связи. Формула справедлива, если используемый код не исправляет ошибок. Применяя код, исправляющий ошибки,

необходимо определить вероятность появления ошибки определенной кратности для данной модели помех. Тогда вероятность возникновения ошибки кратности больше $i$ для кода длиной $n$ [9]

$$p(>i) = \frac{p}{q_0} \sum_{l=\frac{i}{q_0}+1}^{l_{\max}} \left[1 + \frac{n - \left(\frac{2i}{q_0} + 1\right)}{l}\right] \frac{lp_l}{\sum_{l=1}^{l_{\max}} lp_l},$$

а вероятность прохождения сообщений при использовании кода, исправляющего $s_и$ ошибок:

$$P_{прs} = 1 - \frac{p}{q_0} \sum_{l=\frac{s_и}{q_0}+1}^{l_{\max}} \left[1 + \frac{\left[n - \left(\frac{2s_и}{q_0} + 1\right)\right]}{l}\right] \frac{lp_l}{\sum_{l=1}^{l_{\max}} lp_l}.$$

Отсюда нетрудно установить целесообразность использования кодов, исправляющих $s_и$ ошибок, если воспользоваться понятием эквивалентной вероятности ошибки, рассмотренным выше. Названный код имеет смысл применять, когда эквивалентная вероятность ошибки в коде с избыточностью меньше, чем в коде без избыточности. Получение конкретных соотношений и рекомендаций возможно при условии, что известна модель ошибок в канале связи. Обычно при переходе к модели пакетных ошибок требуются коды с высокой корректирующей способностью.

Оценим границы эффективного применения различных видов избыточности в аппаратуре. Предположив, что неисправности в аппаратуре и их последствия в виде отказов возникают во времени в соответствии с распределением Пуассона, используем для оценки последствий возникновения неисправностей пуассоновскую модель ошибок при условии, что в аппаратуре выделены общие узлы с интенсивностью отказов $\lambda_0$, которые не охвачены избыточностью, а также часть аппаратуры с интенсивностью отказов $n\lambda_1$, определяющая собой $n$ индивидуальных цепей формирования и выявления элементов кода, причем выделенная часть аппаратуры охватывается избыточностью [15].

При исправлении последствий неисправностей, т. е. ошибок отказов, необходимо так строить декодирующее устройство, чтобы оно было способно исправлять ошибки, возникающие в отдельных частях аппаратуры. Поток ошибок, который фиксируется приемными устройствами декодирования, включает ошибки, возникающие в кодирующем устройстве и в схеме выявления элементов кода, т. е. на приемной стороне, если канал связи является идеальным и помехи в нем отсутствуют. Полагая, что эквивалентное распределение потока ошибок от отказов аппаратуры кодирующих и декодирующих устройств соответствует распределению Пуассона, нетрудно оценить вероят-

ностные показатели, определяющие качество функционирования технических средств. Вероятность правильного формирования обнаруживающего кода, или вероятность отсутствия ошибок в сформированном коде при условии идеальной помехоустойчивости

$$P_{\text{пер}}(t) = p(j=0) = \exp(-\lambda_{\text{пер}0} - n\lambda_{\text{пер}1})\, t = \exp(-\lambda_{\text{пер}} t),$$

где $\lambda_{\text{пер}}$ — суммарная интенсивность отказов передающей части системы. Если реализуется код, исправляющий одну ошибку отказа, то

$$P_{\text{пер}}(t) = \sum_{j=0}^{1} p(j) = \exp(-\lambda_{\text{пер}0} t) \sum_{j=0}^{1} [(n\lambda_{\text{пер}1} t)^j / j!] \times \\ \times \exp(-n\lambda_{\text{пер}1} t).$$

Как видно из последней формулы, выделены узлы с интенсивностью отказов $\lambda_{\text{пер}0}$, т. е. общие узлы в аппаратуре передающей части, которые не охвачены избыточностью, и $n$ индивидуальных цепей формирования элементов кода с интенсивностью отказов $\lambda_{\text{пер}1}$. Распределение ошибок в формируемой кодовой комбинации для данных цепей соответствует закону Пуассона. В приемной части аналогично можно выделить узлы с интенсивностью отказов $\lambda_{\text{пр}0}$ и $n$ цепей $\lambda_{\text{пр}1}$. Тогда при выявлении кода, исправляющего одну ошибку, вероятность правильного выявления

$$P_{\text{пр}}(t) = \exp(-\lambda_{\text{пр}0} t) \sum_{j=0}^{1} [(n\lambda_{\text{пр}1} t)^j / j!] \exp(-n\lambda_{\text{пр}1} t).$$

Учитывая работоспособность передающей и приемной частей системы, определим вероятность формирования и выявления кода с числом ошибок, не превышающим корректирующую способность кода:

$$P(t) = P_{\text{пер}}(t)\, P_{\text{пр}}(t).$$

Предполагая, что потоки ошибок отказов подчиняются законам распределения Пуассона, получаем композицию двух независимых законов Пуассона: $n(\lambda_{\text{пер}1} t + \lambda_{\text{пр}1} t) = n\lambda_1 t$, или $\lambda_1 = \lambda_{\text{пер}1} + \lambda_{\text{пр}1}$. Таким образом, если формируется код, исправляющий $s_{\text{а}}$ ошибок, то вероятность правильного приема

$$P_{\text{пр}} = \exp(-\lambda_0 t) \sum_{j=0}^{s_{\text{а}}} [(n\lambda_1 t)^j / j!] \exp(-n\lambda_1 t).$$

Пользуясь данной моделью отказов аппаратуры и полагая помехоустойчивость системы идеальной, можно найти границы эффективного применения различных видов аппаратурной избыточности.

Применяя кодовую аппаратурную избыточность для исправления $s_{\text{а}}$ ошибок отказа, необходимо выбрать кодовое расстояние $d_{\text{а}} \geqslant 2s_{\text{а}} + 1$. При введении этой избыточности в сис-

теме возникают элементы, которые не исправляют собственные ошибки сбоя и отказа. Поэтому, чтобы судить об эффективности введения аппаратурной избыточности, следует учитывать наличие в системе двух частей: неизбыточной (некоторые общие элементы согласования приемной и передающей частей с каналами связи, элементы декодирующего устройства, элементы синхронизации и т. д.) и избыточной (цепи формирования элементов кода или отдельных составляющих кодовой комбинации). Если $P$ — вероятность безотказной работы неизбыточной части устройства, а $p(j)$ — вероятность появления $j$ отказов в избыточной части, где исправляется $s_a$ ошибок, то вероятность правильного формирования и выявления сообщения при идеальной помехоустойчивости определяется как вероятность того, что неизбыточная часть функционирует безотказно, а в избыточной возникает не более чем $s_a$ ошибок отказа, т. е. $P_{пр} = P\sum_{j=0}^{s_a} p(j)$.

Если основой функционирования системы является $K$-ичный код с коэффициентом избыточности $\gamma_s$, то возникает $K-1$ сигнальных признаков. Полагая отказы элементов различных каналов сигнальных признаков взаимно независимыми, из биноминального закона распределения можно получить

$$P(j) = C_{K-1}^{j}[1-\exp(-\gamma_s m\lambda_1 t)]^j[\exp(-\gamma_s m\lambda_1 t)]^{K-1-j}.$$

В методе параллельных сигнальных каналов [29] для исправления $s_a$ ошибок необходимо соблюдать условие $K-1=s_a+1$. Учитывая, что $\Sigma_{j=0}^{K-1} p(j)=1$, получаем

$$P_{пр} = P\sum_{j=0}^{s_a} p(j) = P[1-p(s_a+1)] =$$
$$= P[1-[1-\exp(-\gamma_s m\lambda_1 t)]^{s_a+1}].$$

Определим границы эффективного использования недвоичных кодов для повышения надежности. В случае неизбыточного кода вероятность ошибки

$$P_o = 1-\exp(-m\lambda_1 t),$$

в случае избыточного —

$$P_o = 1-P\{1-[1-\exp(-\gamma_s m\lambda_1 t)]^{s_a+1}\}.$$

Если вводимые дополнительные узлы являются абсолютно надежными, т. е. $P=1$, то

$$\{1-\exp(-[\gamma_s]_{гр}\, m\lambda_1 t)\}^{s_a+1} = 1-\exp(-m\lambda_1 t),$$

$$\exp(-[\gamma_s]_{гр}\, m\lambda_1 t) = 1-\sqrt[s_a+1]{[1-\exp(-m\lambda_1 t)]}$$

и граничное значение коэффициента избыточности

$$[\gamma_s]_{\text{гр}} = (-1/m\lambda_1 t) \ln \left\{1 - \sqrt[s_a+1]{[1 - \exp(-m\lambda_1 t)]}\right\}.$$

Анализ полученного соотношения показывает, что при определенных значениях среднего числа отказов аппаратуры $m\lambda_1 t$ величина $[\gamma_s]_{\text{гр}}$ может стать меньше реального значения коэффициента избыточности, необходимого для построения кода, исправляющего $s_a$ ошибок. Поэтому существуют границы эффективного использования кода, в котором функционирует аппаратура системы. Значительный эффект дает код, исправляющий одну, а также 2 — 3 ошибки.

Для построения избыточных технических средств могут быть использованы и систематические двоичные коды, в которых для отыскания места и характера ошибки имеются специальные логические устройства. Система функционирует тогда в избыточном коде, позволяющем исправлять ошибки отказа. Наличие систематического кода вызывает появление в устройстве избыточных элементов, часть которых корректирует ошибки отказа, а часть — нет. Будем считать, что вероятность безотказной работы неизбыточной дополнительно вводимой части составляет $P$. Число избыточных элементов в коде $k = n - m = m(\gamma_s - 1)$. Тогда вероятность безотказной работы корректора, обеспечивающего исправление ошибок отказа:

$$P_{\text{к}} = \exp(-k\lambda_{\text{изб}} t) = \exp[-m(\gamma_s - 1)\lambda_{\text{изб}} t],$$

где $\lambda_{\text{изб}}$ — интенсивность отказа цепи, соответствующей одному избыточному элементу кода.

Вероятность правильного приема кода можно найти с учетом последовательного соединения устройства, функционирующего в коде, исправляющем $s_a$ ошибок, и декомбинатора, который собственные ошибки отказов не исправляет:

$$P_{\text{пр}} = P_{\text{к}} \sum_{j=0}^{s_a} [(\gamma_s m\lambda_1 t)^j / j!] \exp(-\gamma_s m\lambda_1 t),$$

$$P_{\text{пр}} = \exp[-m(\gamma_s - 1)\lambda_{\text{изб}} t] \sum_{j=0}^{s_a} [(\gamma_s m\lambda_1 t)^j / j!] \exp(-\gamma_s m\lambda_1 t).$$

Полученное соотношение можно уточнить, если учесть наличие общих узлов системы, выполняющих функции синхронизации, согласования с каналом и т. д., вероятность безотказной работы которых $P_{\text{общ}} = \exp(-Qm\lambda_1 t)$. Тогда окончательно получаем

$$P_{\text{пр}} = \exp\{-Qm\lambda_1 t - m(\gamma_s - 1)\lambda_{\text{изб}} t - \\ - \gamma_s m\lambda_1 t\} \sum_{j=0}^{s_a} [(\gamma_s m\lambda_1 t)^j / j!].$$

Расчеты показывают, что при наличии общих узлов в системе и дополнительных устройств, которые не исправляют собственные ошибки отказа, наиболее целесообразным является использование малой корректирующей способности кода. В частности, весьма эффективной оказывается реализация кодов, исправляющих одиночную ошибку отказа. Переход к корректирующему коду можно рассматривать как некоторое эквивалентное снижение интенсивности отказов аппаратуры. Например, переход к коду $s_a = 1$ можно считать эквивалентным функционированию системы в коде без избыточности, но при среднем числе отказов в цепи, соответствующей одному элементу кода, равном $[\lambda_1]_{пр}$. Тогда при $P_{общ} = 1$

$$P_к = \sum_{j=0}^{1} \frac{(\gamma_1 m \lambda_1 t)^j}{j!} \exp(-\gamma_1 m \lambda_1 t) = \exp(-m [\lambda_1]_{пр} t),$$

$$[\lambda_1]_{пр} t = \frac{d/dm [\exp(-m [\lambda_1]_{пр} t)]}{\exp(-m [\lambda_1]_{пр} t)}.$$

Откуда

$$[\lambda_1]_{пр} t = \gamma_1 \lambda_1 t \left(1 - \frac{1}{1 + \gamma_1 m \lambda_1 t}\right) + (\gamma_1 - 1) \lambda_{изб} t.$$

Если применяется код, исправляющий $s_a$ ошибок, то

$$[\lambda_1]_{пр} t = \gamma_s \lambda_1 t \left[ \frac{1 - \sum_{j=0}^{s_a - 1} \frac{(\gamma_s m \lambda_1 t)^j}{j!}}{\sum_{j=0}^{s_a} \frac{(\gamma_s m \lambda_1 t)^j}{j!}} \right] + (\gamma_s - 1) \lambda_{изб} t.$$

При $[\lambda_1]_{пр} = \lambda_1$ получаем верхний допустимый предел введения аппаратурной избыточности. Тогда максимально допустимое среднее число отказов аппаратуры, не охватываемой избыточностью, составит

$$[(\gamma_s - 1) \lambda_{изб} t]_{доп} = \lambda_1 t - \gamma_s \lambda_1 t \left[ \frac{1 - \sum_{j=0}^{s_a - 1} \frac{(\gamma_s m \lambda_1 t)^j}{j!}}{\sum_{j=0}^{s_a} \frac{(\gamma_s m \lambda_1 t)^j}{j!}} \right].$$

Для кода, исправляющего одну ошибку отказа, получаем

$$[(\gamma_1 - 1) \lambda_{изб} t]_{доп} = \lambda_1 t - \gamma_1 \lambda_1 t \left[1 - \frac{1}{1 + \gamma_1 m \lambda_1 t}\right],$$

и при $\gamma_1 m \lambda_1 t \ll 1$ величина $[(\gamma_1 - 1) \lambda_{изб} t]_{доп} = \lambda_1 t$.

Анализ полученных соотношений показывает, что граничное значение интенсивности отказов неизбыточной части системы зависит в основном от коэффициента избыточности $\gamma_s$ и значения интенсивности отказов одной цепи формирования и выявления элемента кода $\lambda_1$. При $\lambda_{изб} = \lambda_1$ граничное

значение коэффициента избыточности можно определить из уравнения

$$[\gamma_s]_{\text{гр}}\left[2-\frac{\sum_{j=0}^{s_a-1}\frac{([\gamma_s]_{\text{гр}}\, m\lambda_1 t)^j}{j!}}{\sum_{j=0}^{s_a}\frac{([\gamma_s]_{\text{гр}}\, m\lambda_1 t)^j}{j!}}\right]=2.$$

Если граничное значение оказывается меньше $\gamma_s$, определяемого из информационных пределов избыточности, в частности из предела Хэмминга для двоичного кода, то это указывает на нецелесообразность использования кодовой аппаратурной избыточности при данном значении $s_a$.

Следует отметить, что при большом числе информационных элементов в коде значение $\gamma_s$ невелико. Относительно небольшим становится и объем избыточной аппаратуры, не исправляющей собственные ошибки отказа. Предположив, что $(\gamma_s-1)\lambda_{\text{изб}}t=0$, при $[\lambda_1]_{\text{пр}}=\lambda_1$ можно найти граничное значение коэффициента избыточности из условия $[\gamma_1]_{\text{гр}}\left|2-\right.$ $\left.-\frac{1}{[\gamma_s]_{\text{гр}}\, m\lambda_1 t}\right]=2$. Если вычислить $[\gamma_s]_{\text{гр}}$, то оказывается, что при кодах значительной длины расширяются границы эффективного использования аппаратурной кодовой избыточности.

Оценим границы использования аппаратурной избыточности повторения. До введения этой избыточности кодовое расстояние $d_a=1$. При $\alpha$-кратном повторении кодовое расстояние увеличивается до $d_a=\alpha+1$. Исправление ошибок отказов становится возможным, если $d_a\geqslant 3$, при этом структура решающего органа, который объединяет получаемые параллельные каналы, должна быть такова, что число совпадений $\alpha_c<$ $<\alpha+1$, тогда происходит исправление ошибок отказа. Если $\alpha_c=\alpha+1$, то обнаружение ошибок отказа происходит с кратностью $r_a=\alpha$. Вероятность правильной работы системы можно определить, пользуясь биноминальным законом:

$$P_{\text{пр}}=P\sum_{i=\alpha+1-s_a}^{\alpha+1}C_{\alpha+1}^{i}\,[p(0)]^i\,[1-p(0)]^{\alpha+1-i},$$

где $P$ — вероятность безотказной работы решающего органа, $p(0)$ — вероятность безотказной работы неизбыточного устройства. Учитывая, что $p(0)=\exp(-m\lambda_1 t)$, получаем

$$P_{\text{пр}}=P\sum_{i=\alpha+1-s_a}^{\alpha+1}C_{\alpha+1}^{i}\exp(-im\lambda_1 t)\,[1-\exp(-m\lambda_1 t)]^{\alpha+1-i}.$$

Представляет интерес оценить допустимый объем избыточной аппаратуры, при котором целесообразно введение аппаратурной избыточности повторения. При $P=\exp(-\lambda_0 t)$, где

$\lambda_0$ — интенсивность отказов узлов, не охваченных избыточностью, получим

$$P_{\text{пр}} = \exp(-\lambda_0 t) \sum_{i=\alpha+1-s_a}^{\alpha+1} C_{\alpha+1}^{i} \exp(-im\lambda_1 t)\,[1 - \exp(-m\lambda_1 t)]^{\alpha+1-i}.$$

Если перейти к закону распределения Пуассона, то

$$P_{\text{пр}} = \exp[-\lambda_0 t - (\alpha+1) m\lambda_1 t] \sum_{i=0}^{s_a} \frac{[(\alpha+1) m\lambda_1 t]^i}{i!}.$$

Граничное значение среднего числа отказов неизбыточной части аппаратуры можно оценить из условия равенства вероятности безотказной работы неизбыточной системы и системы с аппаратурной избыточностью повторения, т. е.

$$\exp\{-[\lambda_0 t]_{\text{доп}} - (\alpha+1) m\lambda_1 t\} \sum_{i=0}^{s_a} \frac{[(\alpha+1) m\lambda_1 t]^i}{i!} = \exp(-m\lambda_1 t),$$

откуда $[\lambda_0 t]_{\text{доп}} = -\alpha m\lambda_1 t + \ln\left\{\sum_{i=0}^{s_a} \frac{[(\alpha+1) m\lambda_1 t]^i}{i!}\right\}$.

Зная $[\lambda_0 t]_{\text{доп}}$, можно определить допустимый объем аппаратуры, не охватываемый избыточностью, если известна конструктивно-технологическая база, на которой строятся технические средства АСУ.

Весьма эффективным является совместное использование кодовой аппаратурной избыточности и аппаратурной избыточности повторения. В этом случае уменьшается число резервных каналов. Система может функционировать в коде, обнаруживающем ошибку, с избыточностью $\gamma$ при наличии $\alpha$-кратного повторения. Если используется код с числом переходов $d_0 = 2$, то при $\alpha$-кратном повторении кодовое расстояние $d_a = 2(\alpha+1)$, при $\alpha = 1$ возможно исправление одиночной ошибки отказа и обнаружение двух ошибок, вероятность правильного функционирования системы можно найти из условия, что система работает исправно, если в $i$ параллельных каналах информация не исказилась, а в остальных возникли обнаруживаемые ошибки. Тогда

$$P_{\text{пр}} = \sum_{i=1}^{\alpha+1} [p(0)]^i C_{\alpha+1}^{i} P_{\text{обн}}^{\alpha+1-i},$$

где $p(0)$ — вероятность непоявления ошибки в выходном сигнале любого параллельного аппаратурного канала, $P_{\text{обн}}$ — вероятность обнаружения ошибки в выходном сигнале канала.

Если $p(0) = \exp(-\gamma m\lambda_1 t)$, $P_{\text{обн}} = P_{\text{ку}} P'_{\text{обн}}$, где $P_{\text{ку}}$ — вероятность исправной работы контролирующего устройства, выполняющего функцию обнаружителя ошибок, $P'_{\text{обн}}$ — вероятность

появления обнаруживаемой ошибки в устройстве, то при использовании кода с числом переходов $d_0$ имеем

$$P'_{\text{обн}} = \sum_{j=1}^{d_0-1} p(j),$$

где $p(j)$ — вероятность возникновения в коде ошибки кратности $j$, тогда

$$P_{\text{пр}} = \sum_{i=1}^{\alpha+1} \{C_{\alpha+1}^{i} \exp(i\gamma m\lambda_1 t) P_{\text{ку}}^{\alpha+1-i} [\sum_{j=1}^{d_0-1} p(j)]^{\alpha+1-i}\}.$$

При $d_0 = 2$ и $P_{\text{ку}} = 1$

$$P_{\text{обн}} = p(1) = \gamma m\lambda_1 t \exp(-\gamma m\lambda_1 t), \text{ а}$$

$$P_{\text{пр}} = \sum_{i=1}^{\alpha+1} C_{\alpha+1}^{i} (\gamma m\lambda_1 t)^{\alpha+1-i} \exp[-\gamma m\lambda_1 t(\alpha+1)].$$

Из полученного соотношения видно, что наибольший эффект дает переход от неизбыточной системы к функционированию ее в обнаруживающем коде с однократным повторением. Увеличение кратности повторения до $\alpha = 2$ дает незначительное увеличение вероятности правильного функционирования.

Оценим границы эффективности использования общей избыточности, направленной на исправление ошибок, вызванных действием помех в каналах связи и отказов аппаратуры технических средств. В данном случае, как и ранее, существенное значение имеют исходные предпосылки по моделям действующих помех в каналах связи и моделям, характеризующим последовательность отказов аппаратуры во времени.

Будем считать, что поток ошибок в формируемом и принимаемом кодах, вызванный ненадежностью системы, аппроксимируется распределением Пуассона с математическим ожиданием $n\lambda_1 t$, кроме того, существуют общие узлы системы, вероятность безотказной работы которых $P = \exp(-\lambda_0 t)$. Предположим далее, что поток ошибок в канале связи подчиняется закону распределения Пуассона с математическим ожиданием $n\tau f_{\text{со}}$, тогда вероятностные показатели функционирования ПСПИ можно найти при условии, что для компенсирования действия помех и отказов в систему введена избыточность. Следует различать раздельное и совместное введение избыточности по помехам и отказам. При раздельном введении избыточности можно использовать методику расчета и оценки границ, изложенную выше. При общем введении избыточности, компенсирующей возмущающие действия помех и отказов, необходим совместный расчет помехоустойчивости и надежности. В простейшем случае вероятность появления $i$ ошибок в коде можно найти, полагая, что каждый из случайных процессов подчиняется во времени закону Пуассона. Тогда

$$p(i) = \sum_{j=0}^{i} p(j) p(i-j),$$

где $j$ — число ошибок в коде, возникающих из-за ненадежности аппаратуры, $i-j$ — число ошибок, возникающих под действием помех в линии связи. Учитывая, что

$$p(j)=\frac{(n\lambda_1 t)^j}{j!}\exp(-n\lambda_1 t),\ p(i-j)=\frac{(n\tau f_{co})^{(i-j)}}{(i-j)!}\exp(-n\tau f_{co}),$$

получим $p(i)=\sum_{j=0}^{i}\frac{(n\lambda_1 t)^j(n\tau f_{co})^{i-j}}{j!(i-j)!}\exp[-n(\lambda_1 t+\tau f_{co})]$.

Так как $\sum_{j=0}^{i}\frac{(n\lambda_1 t)^j(n\tau f_{co})^{i-j}}{j!(i-j)!}i!=(n\lambda_1 t+n\tau f_{co})^i$, то

$$p(i)=[\exp(-n\lambda_1 t-n\tau f_{co})]\frac{(n\lambda_1 t+n\tau f_{co})^i}{i!}.$$

Таким образом, при указанных предположениях вероятность появления ошибок в коде подчиняется закону распределения Пуассона с математическим ожиданием $n(\lambda_1 t+\tau f_{co})$. Определим вероятность правильного приема сообщения, когда используется общая кодовая избыточность, рассчитанная на исправление $s$ ошибок, и вероятность безотказной работы общих узлов в системе равна $P$:

$$P_{пр}=P\sum_{i=0}^{s}p(i).$$

С введением избыточности доля общих узлов в системе увеличивается, поэтому $P$ уменьшается, что отрицательно влияет на вероятность правильного приема сообщения, т. е. существуют определенные границы эффективного использования общей избыточности при увеличивающейся сложности аппаратуры с ростом числа исправляемых ошибок $s$. Применим двоичный систематический код с коэффициентом избыточности $\gamma_s$, исправляющий $s$ ошибок, где $s=s_a=s_и$, тогда

$$p(i)=\{[\gamma_s m(\lambda_1 t+\tau f_{co})]^i/i!\}\exp[\gamma_s m(\lambda_1 t+\tau f_{co})],$$

откуда $P_{пр}=\exp[-\lambda_0 t-\gamma_s m(\lambda_1 t+\tau f_{co})]\sum_{i=0}^{s}\{[\gamma_s m(\lambda_1 t+\tau f_{co})]^i/i!\}$.

Определим границы эффективного использования общей кодовой избыточности из условия равенства вероятностей правильного приема для избыточной и соответствующей ей неизбыточной систем:

$$\exp[-m(\lambda_1 t+\tau f_{co})]=\exp\{-[\lambda_0 t]_{доп}-\gamma_s m(\lambda_1 t+\tau f_{co})\}\times$$
$$\times\sum_{i=0}^{s}\{[\gamma_s m(\lambda_1 t+\tau f_{co})]^i/i!\}.$$

Отсюда максимально допустимое среднее число отказов неизбыточной части системы

$$[\lambda_0 t]_{\text{доп}} = -(\gamma_s - 1)\, m\lambda_1 t - (\gamma_s - 1)\, m\tau f_{\text{со}} + \\ + \ln \textstyle\sum_{i=0}^{s} \{[\gamma_s m (\lambda_1 t + \tau f_{\text{со}})]^i / i!\}.$$

При увеличении $m\lambda_1 t$, $m\tau f_{\text{со}}$ возможно большее значение $[\lambda_0 t]_{\text{доп}}$.

В отдельных случаях при передаче информации, особенно на нижних уровнях ПСПИ, могут быть использованы недвоичные, например, полярные, коды, реализуемые на приемной стороне по методу параллельных каналов [29]. Тогда на передающей и приемной сторонах возникает $K-1$ параллельных каналов сигнальных признаков, и вероятность правильного приема составит

$$P_{\text{пр}} = P \textstyle\sum_{i=0}^{s} p_k(i),$$

где $p_{\text{к}}(i)$ — вероятность выхода из строя $i$ каналов сигнальных признаков. Используя биноминальное распределение, получаем

$$p_{\text{к}}(i) = C^i_{K-1} [1 - \exp(-n\lambda_1 t - n\tau f_{\text{со}})]^i [\exp(-n\lambda_1 t - \\ - n\tau f_{\text{со}})]^{K-1-i}.$$

Учитывая, что $K - 1 = s + 1$, $\sum_{i=0}^{K-1} p_{\text{к}}(i) = 1$, окончательно имеем

$$P_{\text{пр}} = P\{1 - [1 - \exp(-n\lambda_1 t - n\tau f_{\text{со}})]^{s+1}\}.$$

Границу эффективного использования недвоичных кодов как средства введения общей кодовой избыточности для повышения надежности и помехоустойчивости можно найти из условия равенства вероятностей ошибки для неизбыточной и соответствующей ей избыточной систем. Тогда для неизбыточного кода вероятность ошибки

$$P_{\text{о}} = 1 - \exp[-m(\lambda_1 t + \tau f_{\text{со}})],$$

для недвоичного избыточного кода

$$P_{\text{о}} = 1 - P\{1 - [1 - \exp(-n\lambda_1 t - n\tau f_{\text{со}})]^{s+1}\}.$$

Полагая $P = 1$, находим граничное максимально возможное значение коэффициента избыточности:

$$[\gamma_s]_{\text{гр}} = -\ln\{1 - \sqrt[s+1]{1 - \exp(-m\lambda_1 t - m\tau f_{\text{со}})}\} / m(\lambda_1 t + \tau f_{\text{со}}).$$

Вычисления показывают, что коды с малой избыточностью имеют широкие пределы использования.

Определим границы применения общей избыточности повторения для повышения помехоустойчивости и надежности. В случае общей избыточности повторения число передач одного и того же сообщения связано с числом резервных цепей, возникающих на приемной и передающей сторонах, соотношением $d = \alpha + 1$. Для обнаружения или исправления ошибок на приемной стороне предусматривается решающий орган, в зависимости от построения которого удается либо обнаружить, либо обнаружить и исправить ошибки. Если для передачи информации в качестве исходного используется неизбыточный код длиной $m$, то при введении общей избыточности повторения по каждому элементу кода возникает $\alpha + 1$ параллельных каналов формирования и выявления. Вероятность правильного приема сообщения найдем из условия, что для правильного функционирования системы является необходимым отсутствие отказов в общих узлах системы и возможно наличие ошибок с кратностью, не превышающей $s$ в принимаемом коде. С учетом действия помех и отказов вероятность искажения одного элемента кода $p = p_{\text{пом}}(1 - p_{\text{отк}}) + p_{\text{отк}}(1 - p_{\text{пом}}) + p_{\text{пом}}p_{\text{отк}}$, где $p_{\text{пом}} = 1 - \exp(-\tau f_{\text{со}})$, $p_{\text{отк}} = 1 - \exp(-\lambda_1 t)$. Вероятность правильного приема элемента кода с повторением

$$P_{\text{пр1}} = \sum_{i=0}^{s} C_{\alpha+1}^{i} p^i (1-p)^{\alpha+1-i}.$$

Вероятность правильного приема сообщения, отображаемого $m$-элементным неизбыточным кодом:

$$P_{\text{пр}} = PP_{\text{пр1}}^{m} = P\left[\sum_{i=0}^{s} C_{\alpha+1}^{i} p^i (1-p)^{\alpha+1-i}\right]^m.$$

Из последнего соотношения следует, что с увеличением корректирующей способности кода $s$ вероятность правильного приема сообщения возрастает. Если используется общая избыточность повторения, то существует некоторое допустимое значение $P$, определяемое для максимально возможного объема аппаратуры, не охваченной избыточностью. Полагая, что $P = \exp(-\lambda_0 t)$, можно при малых $\lambda_1 t$, $\tau f_{\text{со}}$ перейти к распределению Пуассона. Учитывая, что $p_{\text{пом}} \approx \tau f_{\text{со}}$, $p_{\text{отк}} \approx \lambda_1 t$, получаем $p \approx \tau f_{\text{со}} + \lambda_1 t$,

$$C_{\alpha+1}^{i}(\tau f_{\text{со}} + \lambda_1 t)^i (1 - \tau f_{\text{сп}} - \lambda_1 t)^{\alpha+1-i} =$$
$$= \{[(\alpha+1)(\tau f_{\text{со}} + \lambda_1 t)]^i / i!\} \exp[-(\alpha+1)(\tau f_{\text{со}} + \lambda_1 t)].$$

Тогда $P_{\text{пр}} = \exp[-\lambda_0 t - m(\alpha+1)(\tau f_{\text{со}} + \lambda_1 t)] \{\sum_{i=0}^{s} \{(\alpha + 1)(\lambda_1 t + \tau f_{\text{сп}})]^i / i!\}\}^m$.

Определим максимально допустимое среднее число отказов дополнительно вводимой аппаратуры, не охваченной избы-

точностью, при равенстве вероятностей правильного приема сообщения в неизбыточной и соответствующей ей избыточной системах:

$$\exp[-m(\lambda_1 t+\tau f_{co})]=\exp\{-[\lambda_0 t]_{доп}-m(\alpha+1)(\tau f_{co}+\\+\lambda_1 t)\}\{\textstyle\sum_{i=0}^{s}\{(\alpha+1)(\lambda_1 t+\tau f_{co})]^i/i!\}\}^m,$$

откуда

$$[\lambda_0 t]_{доп}=-\alpha m(\lambda_1 t+\tau f_{co})+m\ln\{\textstyle\sum_{i=0}^{s}\{[(\alpha+\\+1)(\lambda_1 t+\tau f_{co})]^i/i!\}\}.$$

С увеличением $s$ значение $[\lambda_0 t]_{доп}$ возрастает, что подтверждает возможность практической реализации общей избыточности повторения. По величине $[\lambda_0 t]_{доп}$, зная конструктивно-технологическую базу технических средств, используемых для построения АСУ, можно определить максимально допустимый объем дополнительной аппаратуры, которая вводится в систему и избыточностью не обладает. Поэтому нетрудно оценить границы эффективного использования и различных совместных сочетаний общей, аппаратурной и информационной избыточности, при которых $s_a \neq s_и \neq s$. Естественно, что получаемые при этом границы определяются принятыми исходными моделями ошибок в каналах связи и отказов в аппаратуре системы. Эти границы с изменением модели ошибок меняются, и поэтому представляет интерес оценить границы использования избыточности для каналов, в которых наблюдается группирование ошибок в пакеты при неидеальной надежности аппаратуры.

Если система строится с кодовой информационной избыточностью, то вероятность правильного приема с учетом пакетного характера ошибок в канале связи можно определить из соотношений [9]

$$P_{пр}=[\exp(-\lambda_0 t-\gamma_s m\gamma_1 t)]\left\{1-\frac{p}{q_0}\sum_{l=\frac{s_и}{q_0}+1}^{l_{max}}\left[1+\right.\right.\\\left.\left.+\frac{n-\left(\frac{2s_и}{q_0}+1\right)}{l}\right]\frac{lp_l}{\sum_{l=1}^{l_{max}}lp_l}\right\}.$$

Для нахождения границ эффективного использования кодовой информационной избыточности следует приравнять вероятности правильного приема сообщения для случая избыточной и неизбыточной передач.

Анализ показывает, что за счет усложнения аппаратуры при переходе к избыточному коду возникают более жесткие границы применения кодовой информационной избыточности, чем это предполагалось в случае идеальной надежности. Значительное влияние оказывает объем дополнительно вводимой аппаратуры,

которая избыточностью не охвачена, и увеличение ее объема вызвано лишь необходимостью декодирования кодов, исправляющих ошибки.

Рассмотрим целесообразность применения различных видов информационной и аппаратурной избыточности, если избыточность вводится отдельно для компенсации ошибок помех и для ошибок отказов. Тогда если используется информационная избыточность, рассчитанная на исправление $s_и$ ошибок, и аппаратурная кодовая избыточность, рассчитанная на исправление $s_а$ ошибок, то вероятность правильного приема

$$P_{пр} = \left\{1 - \frac{p}{q_0} \sum_{l=\frac{s_и}{q_0}+1}^{l_{max}} \left[1 + \frac{n - \left(\frac{2s_и}{q_0} + 1\right)}{l}\right] \times \right.$$

$$\left. \times \frac{lp_l}{\sum_{l=1}^{l_{max}} l\, p_l} \right\} \exp\left(-\lambda_0 t - \gamma_s m \lambda_1 t\right) \sum_{i=0}^{s_а} \frac{(\gamma_s m \lambda_1 t)^i}{i!}.$$

Следует отметить, что независимо от модели ошибок в каналах связи границы использования аппаратурной избыточности при сохранении закона распределения во времени остаются справедливыми и определяются условиями введения избыточности в систему. Однако для исправления последствий действия помех, приводящих к пакетному распределению ошибок в канале, целесообразным является применение кодов с высокой корректирующей способностью, а поэтому численные расчеты в данном случае должны проводиться для больших значений $s_и$, где границы их применения оказываются весьма узкими. В большинстве случаев необходимо оценить функциональную зависимость объема аппаратуры кодирования и декодирования от длины кода и способа декодирования.

Таким образом, конкретные границы использования различного типа избыточности могут быть установлены тогда, когда известны длина кода, модель ошибок в канале связи и закон распределения отказов аппаратуры во времени. Все это можно осуществить при известной структуре кодирующих и декодирующих устройств.

Введение информационной и аппаратурной избыточности позволяет в общем случае снизить вероятность потерь $P_{пот\ н}$. Однако остается доля вероятности потерь, связанных с отказом в обслуживании, а поэтому существенное значение имеет еще одна разновидность избыточности, а именно избыточность в обслуживании. Она позволяет уменьшить время обслуживания информационных потоков, что для АСУ имеет исключительно важное значение, поскольку в задачах, связанных с оперативным управлением, регулированием или контролем, существенную роль играет старение информации, а поэтому задержки,

вызванные длительным хранением информации в запоминающих устройствах, могут привести к потере информации в целом. Вместе с тем перегрузка запоминающих устройств, расположенных в узлах концентрации информации или в приемных частях аппаратуры ПСПИ, вызывает отказы в приеме, и при отсутствии определенных объемов запоминающих устройств информация теряется, что учитывается составляющей потерь с вероятностью $P_{\text{пот о}}$. Вычисление вероятности может быть выполнено, если известна модель входного потока сообщений. При этом различают модель случайного потока и модель регулярных входных потоков сообщений.

Значительную сложность представляет выбор между избыточностью в информации и избыточностью в обслуживании. Как указывалось ранее, введение избыточности в передаваемую информацию увеличивает время обслуживания каждой заявки. В соответствии с этим для компенсации потерь информации, вызванных отказом в обслуживании, необходимо ввести избыточность в обслуживающую систему, что может проявляться в увеличении объема запоминающих устройств (избыточность по памяти) или числа обслуживающих (декодирующих) устройств. Это приводит к увеличению объема аппаратуры системы и тем самым к снижению надежности. Поэтому серьезной задачей при проектировании технических средств сбора и передачи информации в АСУ является определение оптимальных соотношений по всем видам избыточности с учетом возможности обслуживания пересекающихся информационных потоков. При идеальной надежности, когда многоканальное обслуживание не осуществляется, можно рекомендовать в качестве оптимального варианта введение информационной избыточности в виде обнаруживающих кодов с использованием сигналов обратной связи. При многоканальном обслуживании эффективен переход к корректирующим кодам в канале связи. Это приводит к значительному сокращению времени обслуживания каждого сообщения и тем самым вызывает снижение вероятности потерь по причине отказа в обслуживании. Поэтому комплексный информационный подход к оценке потерь информации с учетом всех сторон функционирования технических средств позволяет получить совершенно новые рекомендации, которые в большей степени соответствуют реальным условиям сбора, передачи и обработки информации в АСУ.

## УКАЗАТЕЛЬ ЛИТЕРАТУРЫ


1. Аверин В. И. и др. Выбор комплекса технических средств АСУП. М., 1973. 208 с.
2. Аветов Ю. В., Советов Б. Я., Тимошинов П. М. Каналы связи телемеханики. Л., 1971. 128 с.
3. Бородин Л. Ф. Введение в теорию помехоустойчивого кодирования. М., 1968.
4. Виноград С., Коуэн Дж. Д. Надежные вычисления при наличии шумов. (Пер. с англ.). М., 1968. 112 с.
5. Галлагер Р. Теория информации и надежная связь. (Пер. с англ.). М., 1974. 720 с.
6. Железнов Н. А. Синтез информационных систем и использование избыточности.— В кн.: Использование избыточности в информационных системах. Л., 1970.
7. Клюев Н. И. Информационные основы передачи сообщений. М., 1966. 360 с.
8. Мамиконов А. Г. Методы разработки автоматизированных систем управления. М., 1973. 336 с.
9. Мельников Ю. Н. Достоверность информации в сложных системах. М., 1973. 192 с.
10. Мизин И. А., Уринсон Л. С., Храмешин Г. К. Передача информации в сетях с коммутацией сообщений. М., 1972. 320 с.
11. Справочник проектировщика систем автоматизации управления производством. Ред. Г. Л. Смилянский. М., 1971. 424 с.
12. Хартли Р. Передача информации.— В кн.: Теория информации и ее приложения. Ред. А. А. Харкевич. М., 1959, с. 5—36.
13. Самойленко С. И. Помехоустойчивое кодирование. М., 1966. 240 с.
14. Советов Б. Я. Большие системы сбора и передачи информации. Л., 1970. 118 с.
15. Советов Б. Я. Эффективность введения избыточности в системы передачи телемеханической информации. Л., 1972. 132 с.
16. Советов Б. Я. Теория информации. Ч. 1. Л., 1974. 74 с.
17. Советов Б. Я. Основы построения АСУ. Л., 1975. 134 с.
18. Солодов А. В. Теория информации и ее применение к задачам автоматического управления и контроля. М., 1967. 432 с.
19. Тарасенко Ф. П. Введение в курс теории информации. Томск, 1963. 240 с.
20. Темников Ф. Е., Афонин В. А., Дмитриев В. И. Теоретические основы информационной техники. М., 1971. 424 с.

21. Фано Р. Передача информации. Статистическая теория связи. (Пер. с англ.). М., 1965. 438 с.
22. Финк Л. М. Теория передачи дискретных сообщений. М., 1963. 576 с.
23. Харкевич А. А. Борьба с помехами. М., 1963. 276 с.
24. Харкевич А. А. Очерки общей теории связи. М.. 1955. 268 с.
25. Харкевич А. А. Основы радиотехники. М., 1962. 558 с.
26. Четвериков В. Н. Преобразование и передача информации в АСУ. М., 1974. 320 с.
27. Шастова Г. А. Кодирование и помехоустойчивость передачи телемеханической информации. М., 1966. 454 с.
28. Шеннон К. Работы по теории информации и кибернетике. (Пер. с англ.). М., 1965. 438 с.
29. Юргенсон Р. И. Помехоустойчивость цифровых систем передачи телемеханической информации. Л., 1971.
30. Юргенсон Р. И. Передача информации и методы борьбы с помехами в телемеханике. Ч. 1. Л., 1968.

## ОГЛАВЛЕНИЕ


Предисловие . . . . . 3
Введение . . . . . 5
§ 1. Автоматизированное управление производством . . —
§ 2. Информационный характер процесса принятия решения . . . . . 7
§ 3. Основные подсистемы комплекса технических средств АСУ . . . . . 10
§ 4. Место и роль теории информации при разработке АСУ . . . . . 12

Глава I. **Информационный процесс в АСУ** . . . . . 16

§ 1. Фазы преобразования информации в АСУ . . . . —
§ 2. Понятие сообщения и кода . . . . . 19
§ 3. Кодирование информации в АСУ . . . . . 24
§ 4. Информационные сети и информационные потоки . . 30
§ 5. Виды модуляции и их реализация в информационных сетях АСУ . . . . . 35

Глава II. **Количественные характеристики информации** . . . . . 43

§ 1. Количество информации в равновероятных сообщениях . . . . . —
§ 2. Количество информации в неравновероятных сообщениях . . . . . 46
§ 3. Понятие и свойства энтропии . . . . . 49
§ 4. Энтропия источника и энтропия сообщения . . . . 55
§ 5. Связь количества информации с эффективностью управления . . . . . 58

Глава III. **Скорость передачи информации и пропускная способность каналов связи АСУ** . . . . . 61

§ 1. Понятие канала связи. Характеристики каналов связи АСУ . . . . . —
§ 2. Простейшая структурная схема формирования, передачи и выявления сообщений . . . . . 67

§ 3. Скорость передачи информации и пропускная способность канала связи . . . . . 73
§ 4. Методы приближения скорости передачи информации к пропускной способности дискретного канала связи 77
§ 5. Скорость передачи информации и пропускная способность непрерывного канала связи . . . . . 81

Глава IV. **Дискретная информация в АСУ** . . . . . 84

§ 1. Кодирование дискретной информации в каналах связи без шума . . . . . —
§ 2. Методы построения оптимальных кодов . . . . . 86
§ 3. Кодирование дискретной информации в каналах связи с шумом . . . . . 90
§ 4. Информационные пределы избыточности . . . . . 96
§ 5. Информационные процессы при кодировании и декодировании . . . . . 98
§ 6. Граничное значение вероятности ошибки при производительности источника, превышающей пропускную способность канала связи . . . . . 104
§ 7. Эквивалентная вероятность ошибки . . . . . 105

Глава V. **Непрерывная информация в АСУ** . . . . . 108

§ 1. Дискретизация непрерывной информации . . . . . 109
§ 2. Энтропия непрерывного дискретизированного процесса . . . . . 113
§ 3. Скорость передачи непрерывной дискретизированной информации и пропускная способность канала связи 116
§ 4. Погрешности квантования непрерывной функции по времени и по амплитуде . . . . . 118
§ 5. Оптимальное квантование непрерывной функции по амплитуде . . . . . 121

Глава VI. **Выделение сигнала на фоне помех** . . . . . 125

§ 1. Геометрическое представление сигнала в пространстве —
§ 2. Выделение детерминированного сигнала на фоне нормального шума . . . . . 128
§ 3. Выделение случайного сигнала на фоне нормального шума . . . . . 132
§ 4. Методы приема сигналов . . . . . 134
§ 5. Выделение шумоподобных сигналов на фоне помех 139

Глава VII. **Информационный подход к оценке качества функционирования комплекса технических средств АСУ** . . . . . 143

§ 1. Основные критерии качества функционирования комплекса технических средств АСУ . . . . . —
§ 2. Информационный подход к оценке верности преобразования информации . . . . . 146
§ 3. Избыточность как средство обеспечения заданного качества функционирования технических средств АСУ 156

Указатель литературы . . . . . 181

## ОПЕЧАТКИ

| Страница | Строка | Напечатано | Следует читать |
|---|---|---|---|
| 101 | 4-я сверху | $P(e/y_{0i}) = = \sum_{i=1}^{M} p(x_{0j})\, p(y_{0i}/x_{0j})$ при $i \neq j$. | $P(e/y_{0i}) = = \sum_{j=1}^{M} p(x_{0j})\, p(y_{0i}/x_{0j})$ при $i \neq j$. |
| 171 | 13-я сверху | $P_к = \sum_{j=0}^{1} \frac{(\gamma_1 m \lambda_1 t)^j}{j!} \exp \times$ $\times (-\gamma_1 m \lambda_1 t) =$ | $P_к \sum_{j=0}^{1} \frac{(\gamma_1 m \lambda_1 t)^j}{j!} \exp \times$ $\times (-\gamma_1 m \lambda_1 t) =$ |

Зак. № 931

53 коп.

ИЗДАТЕЛЬСТВО ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА 1977

53 коп.

ИЗДАТЕЛЬСТВО ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА 1977

Б. Я. СОВЕТОВ ТЕОРИЯ ИНФОРМАЦИИ